

AD _____

Award Number: W81XWH-05-1-0041

TITLE: Prostate Dose Escalation by a Innovative Inverse Planning-Driven IMRT

PRINCIPAL INVESTIGATOR: Lei Xing, Ph.D

CONTRACTING ORGANIZATION: Stanford University
Stanford, CA 94305-5401

REPORT DATE: November 2008

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE 1 Nov 2008		2. REPORT TYPE Final		3. DATES COVERED 1 Nov 2004 – 31 Oct 2008	
4. TITLE AND SUBTITLE Prostate Dose Escalation by a Innovative Inverse Planning-Driven IMRT				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-05-1-0041	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Lei Xing, Ph.D E-Mail: lei@reves.stanford.edu				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Stanford University Stanford, CA 94305-5401				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT The goal of this project is to develop innovative inverse planning techniques for prostate radiation therapy. In the last funding period significant progress has been made toward the goal of the project. In the past funding year, the PI group has continued to advance prostate radiation therapy techniques. A few important milestones have been achieved toward the goal of the project. These include: (i) Proposed a novel total-variation regularization based inverse planning technique for prostate IMRT; (ii) Applied compressed sensing concept to IMRT and IMAT inverse planning and established a multi-objective optimization framework; (iii) Developed a voxel-specific penalty scheme for TRV-based inverse planning; (iv) Established a cine-EPID image retrospective dose reconstruction in IMRT dose delivery for adaptive planning and IMRT dose verification. These works are both timely and important and should lead to widespread impact on prostate cancer management.					
15. SUBJECT TERMS Prostate Cancer					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	155	19b. TELEPHONE NUMBER (include area code)

Table of Contents

Introduction.....	4
Body	4
Key Research Accomplishments	8
Reportable Outcomes	8
Conclusions.....	10
References.....	11
Appendices.....	12

I. INTRODUCTION

This Idea Award (PC040282/W81XWH05-1-0041), entitled “Prostate Dose Escalation by Innovative Inverse Planning-Driven IMRT”) was awarded to the principal investigator (PI) for the period of Nov 1, 2004—Oct. 31, 2007. Due to a delay in having a postdoctoral fellow with right training background on board during the course of the project, an one year no-cost-extension was filed and granted by DOD. Thus the final completion date of the project is Oct. 31, 2008. This is the final report of the project.

The goal of this project is to improve current prostate IMRT by establishing a novel inverse planning framework. Under the generous support from the U.S. Army Medical Research and Materiel Command, the PI’s research team has made significant progress toward the general goal of the project and contributed greatly to prostate cancer research. A number of highly significant conference abstracts and refereed papers have been resulted from the support. The preliminary data obtained under the support of the grant has also enabled the PI to start new research initiatives. In this report, the research activities and accomplishments of the PI are highlighted.

II. RESEARCH AND ACCOMPLISHMENTS

We have continued to improve the prostate IMRT dose optimization techniques and made significant progress in treatment planning for image guided prostate radiation therapy and related issues. We proposed, for the first time, a total-variation regularization (TVR) method for improved IMRT inverse planning¹ (Appendix 10). An emerging signal processing technique, compressed sensing, has been applied for IMRT inverse planning² (Appendix 12). A voxel-specific penalty scheme with TVR framework has been investigated (Appendix 11). The new planning scheme has been evaluated using a large number of archived prostate cases and the results are very promising. These works lay foundation for the next generation of inverse planning techniques and may significantly impact the clinical prostate IMRT. In addition, we developed an effective inverse planning algorithm for intensity modulated arc therapy (IMAT)³⁻⁶ with incorporation of beam eye’s view dosimetrics⁷, which improves the computational speed by an order of magnitude as compared with the existing dose optimization technique while improving the final dose distribution. The study should have widespread impact on clinical IMRT, IMAT and image guided radiation therapy (IGRT) in the future.

Using TVR for IMRT inverse planning with field specific numbers of segments: Traditional inverse planning algorithms for step-and-shoot IMRT (SS-IMRT) divide the beam's eye view of the planning target volume (PTV) into small beamlets^{8, 9}. The beamlet intensities are optimized as a linear least square solution, and a leaf sequencing algorithm is then applied on the resultant beam intensity map to generate a set of deliverable beam segments¹⁰⁻¹⁴. Since the physical constraints of the MLC are not included, these algorithms

usually result in a treatment plan with a large number of beam apertures. DAO algorithms have been proposed to reduce the number of segments including naturally the MLC hardware constraints in the optimization^{15, 16}. However, because the delivered dose depends on the aperture shapes non-linearly and the optimization problem is non-convex, random search algorithms, such as simulated annealing, are commonly employed. The computation is intensive and requires tuning of multiple algorithm parameters in the searching and cooling schedules. Moreover, the number of segments for each field needs to be determined before the calculation, which reduces the degree of freedom of the decision variable space and compromises the optimality of the final solution. We have recently proposed an efficient TVR-based inverse planning algorithm¹. Instead of directly applying the physical constraints of the apertures, we include a TVR term in the least square optimization to encourage the computed field intensity maps close to be deliverable using a small number of segments. The TVR term calculates the sum of absolute values of the derivatives of the final fluences and force the optimized beam intensity close to be piece-wise constant. This removes the need for the aperture constraint in the DAO approach. The algorithm is able to optimize the number of segments for each field based on the modulation complexity. The proposed algorithm requires only a simple adaptation of the traditional beamlet-based optimization. We have also successfully reformulated the TRV optimization as a quadratic programming problem. The problem is then efficiently solved using the standard quadratic programming. Although not exactly equivalent, the TVR in the optimization implies the aperture constraint on the intensity maps. The resulting optimized intensity maps are more readily achievable using apertures, at the price of slightly degraded

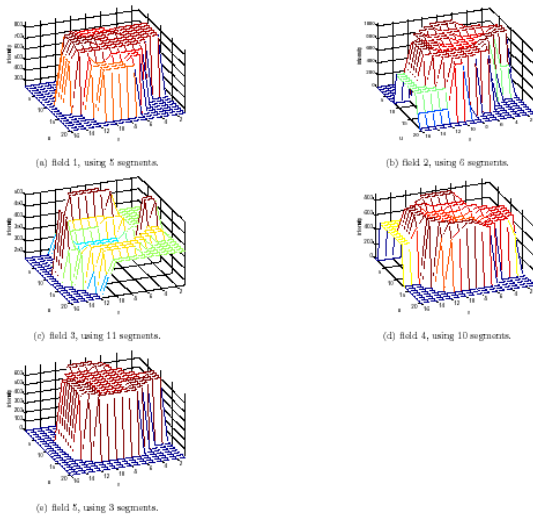


Fig. 1 Fluence maps obtained using the TRV method for a 5-field prostate IMRT case. In this prostate IMRT, five fields were used at angles of 35, 110, 180, 250 and 325 degrees.

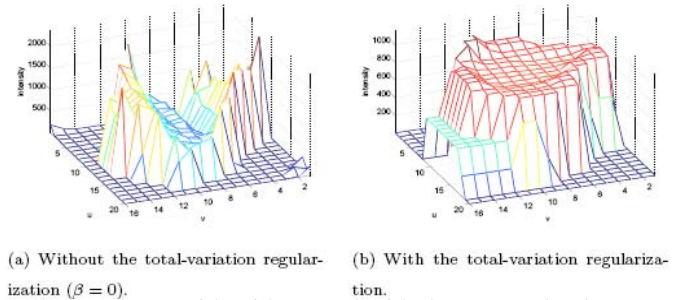


Fig. 2 Fluence maps with without and with the TVR. The data shown in the right is from the 2nd field of Fig. 1

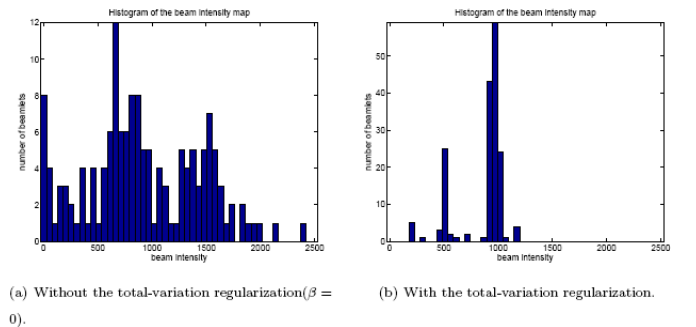


Fig. 3 Histogram of the beam intensities of Fig. 2.

dose distribution as compared to that of the beamlet based plan, but much better than that of DAO for the same number of total segments.

Fig. 1 shows the actual fluence maps obtained using the TVR algorithm for the prostate patient. The resultant plan using our method has different levels of modulation for different fields. It is seen that the

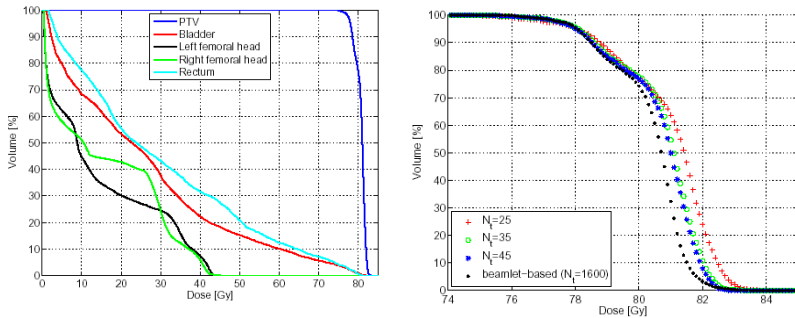


Fig. 4 DVHs of the prostate IMRT plan obtained using the TVR approach (left). The PTV DVH obtained using beamlet-based optimization and the TVR approach with N_t is shown in the right (note the x-axis stars from 74 Gy).

The left of Fig. 4 shows the DVHs when $N_t=35$. The plans are normalized such that 95% of the PTV volume receives 100% prescribed dose (78Gy). The right of Fig. 4 shows the PTV DVHs when different N_t 's are used. As N_t increases, PTV coverage improves. The result using a beamlet-based optimization is also included in the plot. It is seen that, with a relatively small N_t (say $N_t=35$, corresponding to an average of 7 segments per beam), our algorithm achieves a highly conformal dose distribution, which is comparable to the optimal result with an extremely large number of segments. The resultant plan using our method has different levels of modulation for different fields. In this prostate case, the numbers of apertures are 5, 6, 11, 10, and 3, for fields 1 to 5. Using a commercial treatment planning system (Eclipse, Varian Medical Systems), the numbers of apertures for different fields are 11, 12, 13, 12 and 13, respectively. Our algorithm reduces the total number of apertures from 61 to 35. In another example of 7-field IMRT, the optimal results were achieved using very few numbers of apertures, 1, 4, 3, 4, 4, 8, and 4 for fields 1 to 7. Using the Eclipse system, the numbers of apertures for different fields are 17, 16, 13, 15, 18, 14, and 14, respectively. Our algorithm reduces the total number of apertures from 107 to 28. It is important to let the algorithm to optimize the number of apertures for each field based on the field specific modulation complexity.

The proposed algorithm suppresses the dispensable intensity modulation of fluence maps using TVR. The number of delivered segments is significantly reduced without compromising the conformity of the dose distribution. Furthermore, our algorithm optimizes the individual number of apertures for different fields, based on the complexity of the required modulation.

TVR greatly reduces the complexity of the intensity map, and the resulting intensity map is close to a piece-wise constant function. Fig. 2 shows the optimized beamlet intensity of the 2nd field with and without the TVR. This effect is better illustrated in Fig. 3, which plots the histograms of Fig. 2. The resulting intensity map is close to a piece-wise constant function.

Compressed sensing technique as applied to IMRT inverse planning: In IMRT, the treatment plan is selected from a large pool of physically feasible solutions by optimization of an objective function. The solution depends on the choice of objective function and constraints applied to the optimization. An important characteristic that has not been utilized is that the IMRT solution space is highly degenerated in the sense that there are usually a large number of IMRT plans for the same prescription¹⁷. While these plans yield similar dose distributions satisfying the prescription and constraints, the fluence maps of the plans can be dramatically different. Therefore, it is possible to stipulate constraints in the search of the optimal beamlet intensity such that the resultant number of segments is greatly reduced while the dose distribution is not severely deteriorated. Encouraged by the success of TVR method as described above, we reformulated the algorithm in terms of the emerging compressed sensing theory in signal processing². Since an actual fluence map with a small number of segments must

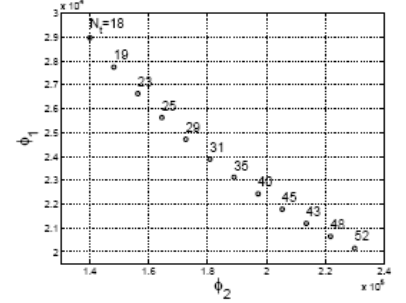


Figure 5 The calculated Pareto frontier of the prostate plan. The derived number of segments (N_t) corresponding to each data point is marked in the plot.

be piece-wise constant and its derivative is sparse, we propose a compressed sensing method to encourage a sparse solution in search of the optimal fluence map and therefore to reduce the total number of segments. The treatment planning is modeled as multi-objective optimization problem with one objective on the dose performance and the other on the sparsity of the solution. A Pareto frontier is calculated, and the achieved dose distributions associated with the Pareto efficient points are evaluated using clinical acceptance criteria. The clinically acceptable dose distribution with the smallest number of segments is chosen as the final solution. The method is demonstrated in the application of fixed-gantry IMRT on a prostate patient and our result shows that the total number of segments is greatly reduced while a satisfactory dose distribution is still achieved. In figure 5 we show the Pareto front plot for a prostate case. As compared to our early TRV work, a practical advantage here is that it effectively eliminate the need for the empirical determination of the TV parameter and makes it possible to obtain truly optimal solution. The method can also be applied in other applications, such as IMAT.

Voxel specific penalty scheme for TRV-based inverse planning: The above TRV-based inverse planning has also been improved by incorporating voxel-specific penalty. In the existing inverse planning algorithms, a structure specific weighting factor and prescription are usually assigned to each structure (target or sensitive structures)^{17, 18}. This type of penalty scheme limits the solution space and often leads to sub-optimal plan (to give a comprehensive example, one can imagine the consequence when two or more structures in a system are restricted to take a single importance). To be able to assess more candidate plans, it is necessary to establish a voxel dependent penalty scheme in which the penalty at a voxel depends not

only on the dose discrepancy but also on the physical and clinical requirements on the point. A natural way to achieve this is to lift the restriction of uniform prescription for a structure so that the prescription at a voxel can be adjusted independently with the goal of finding the optimal final dose distribution. The voxel-specific prescription is simply a set of guidance parameters that help the optimizer to find the optimal solution. We have recently established a framework for this type of TRV-based inverse planning scheme and the manuscript is in preparation (a draft is attached in the Appendix 11). This work further enhances our initiatives on TRV and compressed sensing based IMRT inverse planning and provides an additional means to ensure adequate dose coverage of the prostate while sparing the rectum and bladder.

III. KEY RESEARCH ACCOMPLISHMENTS

- Proposed a novel total-variation regularization based inverse planning technique for prostate IMRT.
- Applied compressed sensing concept to IMRT and IMAT inverse planning and established a multi-objective optimization framework.
- Developed a voxel-specific penalty scheme for TRV-based inverse planning.
- Established a cine-EPID image retrospective dose reconstruction in IMRT dose delivery for adaptive planning and IMRT dose verification.

IV. REPORTABLE OUTCOMES

The following is a list of publications resulted from the grant support in the last funding period.

Refereed publications:

1. Lee L, Le Q, **Xing L**: Retrospective IMRT dose reconstruction based on cone-beam CT and MLC log-file. *International Journal of Radiation Oncology, Biology and Physics*, 70: 634-644, 2008.
2. Chao M, Schreibamnn E, Li T, **Xing L**, Automated contour mapping with a regional deformable model, *International Journal of Radiation Oncology, Biology and Physics*, 70: 599-608, 2008.
3. Thornydy B, Koong A, **Xing L**, Reducing respiratory motion artifacts in radionuclide imaging through retrospective stacking: A simulation study, *Linear Algebra and its Applications*, 428: 1325-1344, 2008.
4. **Xing L**, Quality assurance of PET/CT for radiation therapy, *International Journal of Radiation Oncology, Biology and Physics* 71, ,38-41, 2008.
5. Mao W, Lee L, **Xing L**, Design of multi-purpose phantom and automated software analysis tool for quality assurance of onboard kV/MV imaging system, *Medical Physics* 35, 1497-1506, 2008.
6. Schreibmann E., Thorndyke B, **Xing L**, 4D-4D Image registration for image guided radiation therapy (IGRT), *International Journal of Radiation Oncology, Biology and Physics* 71, 578-586, 2008. - highlighted article of the issue.
7. Wiersma R. and **Xing L**, Real-time monitoring of implanted fiducials using onboard kV and treatment MV beams, *Medical Physics* 35, 1191-1198, 2008.
8. Mao W, Wiersma R, **Xing L**, Fast fiducial detection algorithm for onboard MV and kV imaging systems, *Medical Physics* 35, 1942-1949, 2008.
9. Xie Y, Djajaputra D, King C, Hossain S, Ma L, **Xing L**: Intrafraction motion of prostate in hypofractionated radiation therapy. *International Journal of Radiation Oncology, Biology and Physics* 72, 236-246, 2008.
10. Xie Y, Chao M, Lee P, **Xing L**: Feature-based rectal contour propagation from planning CT to cone beam CT, *Medical Physics* 35, 4450-4459, 2008.
11. Wang J, Li T, Liang Z, **Xing L**, Dose reduction in kV cone beam CT for radiation therapy, *Physics in Medicine and Biology* 53, 2897-2909, 2008.

12. Chao M, Schreibamnn E, Li T, **Xing L**, Contour propagation from planning CT to cone beam CT (CBCT), *Physics in Medicine and Biology* 53, 4533-4542, 2008.
13. Mao W, Riaz N, Lee L, Wiersma R, **Xing L**, A fiducial detection algorithm for real-time image guided IMRT based on simultaneous MV and kV imaging, *Medical Physics* 35, 3554-3564, 2008.
14. Paquin D, Levy D, **Xing L**, Hybrid multistage landmark and deformable image registration, *Medical Physics*, in press, 2008.
15. Wang C, Chao M, Lee L., **Xing L**, MRI-based Treatment Planning with Electron Density Information Mapped from CT Images: A Preliminary Study, *Technology in Cancer Research Treatment* 7, 341-348, 2008.
16. Lee L, Mao W, Xing L, The use of EPID-measured leaf sequence files for IMRT dose reconstruction in adaptive radiation therapy, *Medical Physics* 35, 5019-5029, 2008.
17. Wang J, Li T, **Xing L**: Low-Dose CBCT Imaging for External-Beam Radiotherapy. *Medical Physics*, in press, 2008.
18. Zhu L, Lee L, Ma Y, Ye Y, Mazzeo R, and **Xing L**, Using total-variation regularization for inverse planning with field specific numbers of segments, *Physics in Medicine and Biology* 53, 6653-6672, 2008.
19. Zhu L, Wang J, **Xing L**, Noise Suppression in Scatter Correction for Cone-Beam CT, *Medical Physics*, conditionally accepted, 2008.
20. Wang J, Li T, **Xing L**: Dose reduction in fluoroscopic imaging. *Medical Physics*, accepted, 2008.
21. Xie Y, Chao M, **Xing L**: Modeling the Shear Movement of the Lungs During Respiration Using Tissue Feature-Based Image Registration, *International Journal of Radiation Oncology, Biology and Physics*, conditionally accepted, 2008.
22. Riaz N, Shanker P, Gudmundsson O, Wiersma R, Mao W, Widrow B, and **Xing L**, Predicting respiratory tumor motion with Multi-dimensional Adaptive Filters and Support Vector Regression, *International Journal of Radiation Oncology, Biology, Physics*, conditionally accepted, 2008.
23. Liu W, Wiersma R, Mao W, Luxton G, and **Xing L**, Real-time 3D internal marker tracking during arc radiotherapy by use of combined MV-kV imaging, *Physics in Medicine and Biology* 53, 7197-7213, 2008.
24. Zhu L, and **Xing L**, Search for IMRT solutions with piece-wise constant fluence maps using compressed sensing technique, *Medical Physics*, submitted, 2008.
25. Ma Y, Park S, Suh T, Keall P, **Xing L**, Four-dimensional inverse planning with inclusion of implanted fiducials in IMRT segmented fields, *International Journal of Radiation Oncology, Biology and Physics*, submitted, 2008.
26. Zhu L, Wang J, **Xing L**, Scatter correction in cone beam CT, submitted to *Medical Physics*, 2008.
27. Ma Y, Lee L, **Xing L**, Inverse planning for 4D modulated arc therapy, *Medical Physics*, submitted, 2008.
28. Ma Y, **Xing L**, Incorporation of prior knowledge into treatment planning of modulated arc therapy, *International Journal of Radiation Oncology, Biology, Physics*, submitted, 2008.

Book and Book Chapters

1. Keall P and **Xing L**, Image Guided and Adaptive Therapy, The Textbook of Radiation Oncology, Phillips T and Leibel S (editors), Saunders, 2008.
2. Xia P, Almos H, **Xing L**, Intensity Modulated Radiation Therapy, The Textbook of Radiation Oncology, Phillips T and Leibel S (editors), Saunders, 2008.
3. **Xing L**, Lee L, Timmerman R, Image Guided Adaptive Radiation Therapy, in Image Guided and Adaptive Therapy, Timmerman R. and **Xing L** (editors), *Image Guided and Adaptive Radiation Therapy*, Lippincott, Williams, and Wilkins, Baltimore, MD, in processing (to appear in 2009).
4. **Xing L**, Image Guided Intensity Modulated Radiation Therapy, in Mathematical Methods in Biomedical Imaging and IMRT, Censor Y, Jiang M, Louis A.K (Editors), Springer-Verlag, 2008.
5. Wiersma R, Riaz N, **Xing L**, Real-Time Guided Image Guided Intensity Modulated Radiation Therapy, in Biomedical Imaging and IMRT Inverse Planning, Jiang M, Censor Y, Wang G (Editors), Medical Physics Publishing, 2009.

Conference abstract

1. W Mao, N Riaz, K Lee, R Wiersma, C King, A Hsu, G Luxton, L Xing, Using Treatment Beam Imaging to Monitor Prostate Motion In Near Real-Time On a Conventional LINAC, 2008 Annual Meeting of AAPM, Houston, TX, July, 2008.
2. W Liu, W Mao, R Wiersma, G Luxton, N Riaz, L Xing, Nearly Real-Time Tumor-Position Monitoring During Arc Therapy with Combined MV and KV Imaging, 2008 Annual Meeting of AAPM, Houston, TX, July, 2008.
3. N Riaz, R Wiersma, W Mao, L Xing Prediction of Fiducial Motion in Respiratory Tumors for Image-Guided Radiotherapy, 2008 Annual Meeting of AAPM, Houston, TX, July, 2008.

4. L Zhu, L Xing, Scatter Correction for Cone-Beam CT in Radiation Therapy, 2008 Annual Meeting of AAPM, Houston, TX, July, 2008.
5. Y Xie, M Chao, L Xing, 4D CT Image-Based Modeling the Deformation and Sliding Motion of Lungs, 2008 Annual Meeting of AAPM, Houston, TX, July, 2008.
6. Y Xie, D Djajaputra, C King, S Hossain, L Ma, L Xing, Intrafraction Motion of Prostate in Cyberknife Hypofractionated Radiotherapy, 2008 Annual Meeting of AAPM, Houston, TX, July, 2008.
7. N Riaz, W Mao, R Wiersma, L Xing, Improving Tracking of Implanted Radio-Opaque Markers On MV and KV Imaging with Techniques From Computer Vision and Machine Learning, 2008 Annual Meeting of AAPM, Houston, TX, July, 2008.
8. M Chao, Y Xie, L Xing, Image-Based Modeling of Tumor Shrinkage Or Growth: Towards Adaptive Radiation Therapy of Head-And-Neck Cancer, 2008 Annual Meeting of AAPM, Houston, TX, July, 2008.
9. J Wang, T Li, J Liang, L Xing Dose Reduction in Kilovoltage Cone-Beam Computed Tomography for Radiation Therapy, 2008 Annual Meeting of AAPM, Houston, TX, July, 2008.
10. W Mao, N Riaz, K Lee, R Wiersma, L Xing Tracking Multiple Moving Fiducials During Treatment Based On Simultaneous Onboard KV and Treatment MV Imaging, 2008 Annual Meeting of AAPM, Houston, TX, July, 2008.
11. K Lee, W Mao, L Xing, The Use of EPID-Measured Leaf Sequence Files and On-Treatment Cone-Beam CT for Dose Reconstruction of IMRT Delivery, , 2008 Annual Meeting of AAPM, Houston, TX, July, 2008.
12. L Xing, IMRT Dose Painting Guided by Biological and Physiological Imaging, invited talk in 2008 Annual Meeting of AAPM, Houston, TX, July, 2008.
13. L Xing, 4D Imaging for Radiation Therapy, invited talk in 2008 Annual Meeting of RSNA, Chicago, IL, Dec, 2008.
14. Y Xie, D Djajaputra, C. R. King, S. Hossain, L. Ma, L Xing, Intrafractional Motion of the Prostate During Hypofractionated Radiotherapy, Annual meeting of 2008 ASTRO, Boston, MA, Sept, 2008.
15. J. Wang, T. Li, L. Xing, Low-dose CBCT Imaging for External Beam Radiotherapy, Annual meeting of 2008 ASTRO, Boston, MA, Sept, 2008.
16. N. Riaz, P. Agram, O. Gudmundsson, R. Wiersma, W. Mao, L. Xing, Predicting Fiducial Motion in Respiratory Tumors for Image Guided Radiotherapy, Annual meeting of 2008 ASTRO, Boston, MA, Sept, 2008.
17. R.D. Wiersma, W. Mao, L. Xing, Real-time Tracking of Implanted Fiducial Markers using Combined Onboard kV Fluoroscopy and MV EPID Imaging, Annual meeting of 2008 ASTRO, Boston, MA, Sept, 2008.

US Patent

A disclosure entitled “Radiation therapy inverse treatment planning using a regularization of sparse segments” has been filed to the Office of Technology Licensing (OTL) and a patent application will be filed in a few months (Stanford Docket No. S08-277).

IV. CONCLUSIONS

In the past funding year, the PI group has contributed greatly to advance prostate radiation therapy techniques. A few important milestones have been achieved toward the goal of the project. These include: (i) Proposed a novel total-variation regularization based inverse planning technique for prostate IMRT; (ii) Applied compressed sensing concept to IMRT and IMAT inverse planning and established a multi-objective optimization framework; (iii) Developed a voxel-specific penalty scheme for IMRT inverse planning; (iv) Established a cine-EPID image retrospective dose reconstruction in IMRT dose delivery for adaptive planning and IMRT dose verification. We anticipate that these works will lead to widespread impact in clinical prostate IMRT.

References:

-
1. Zhu, L.; Lee, L.; Ma, Y.; Ye, Y.; Mazzeo, R.; Xing, L., Using total-variation regularization for inverse planning with field specific numbers of segments. *Phys Med Biol* **2008**, 53, 6653–6672.
 2. Zhu, L.; Xing, L., Search for IMRT inverse plans with piece-wise constant fluence maps using compressed sensing techniques. *Medical Physics* **2008**, submitted.
 3. Yu, C. X., Intensity-modulated arc therapy with dynamic multileaf collimation: an alternative to tomotherapy. *Physics in Medicine & Biology* **1995**, 40, (9), 1435-49.
 4. Crooks, S. M.; Wu, X.; Takita, C.; Matzich, M.; Xing, L., Aperture modulated arc therapy. *Physics in Medicine and Biology* **2003**, 48, (10), 1333-1344.
 5. Wong, E.; Chen, J. Z.; Greenland, J., Intensity-modulated arc therapy simplified. *Int J Radiat Oncol Biol Phys* **2002**, 53, (1), 222-35.
 6. Otto, K., Volumetric modulated arc therapy: IMRT in a single gantry arc. *Med Phys* **2008**, 35, (1), 310-7.
 7. Ma, Y.; Lee, L.; Xing, L., Incorporation of prior knowledge into treatment planning of modulated arc therapy. *International Journal of Radiation Oncology, Biology, Physics* **2008**, submitted.
 8. Ezzell, G. A.; Galvin, J. M.; Low, D.; Palta, J. R.; Rosen, I.; Sharpe, M. B.; Xia, P.; Xiao, Y.; Xing, L.; Yu, C. X., Guidance document on delivery, treatment planning, and clinical implementation of IMRT: report of the IMRT Subcommittee of the AAPM Radiation Therapy Committee. *Med Phys* **2003**, 30, (8), 2089-115.
 9. Xing, L.; Wu, Q.; Yang, Y.; Boyer, A., Physics of IMRT. In *Intensity-Modulated Radiation Therapy: A Clinical Perspective*, Mundt, A.; Roeske, J., Eds. BC Decker Inc: Hamilton, London, 2005; pp 20-52.
 10. Bortfeld, T.; Boyer, A. L.; Schlegel, W.; Kahler, D. L.; Waldron, T. J., Realization and verification of three-dimensional conformal radiotherapy with modulated fields. *International Journal of Radiation Oncology, Biology, Physics* **1994**, 30, (4), 899-908.
 11. Bortfeld, T., Optimized planning using physical objectives and constraints. *Seminars in Radiation Oncology* **1999**, 9, (1), 20-34.
 12. Xing, L.; Chen, G. T. Y., Iterative algorithms for Inverse treatment planning. *Physics in Medicine & Biology* **1996**, 41, (2), 2107-23.
 13. Boyer, A. L.; Yu, C. X., Intensity-modulated radiation therapy with dynamic multileaf collimators. *Seminars in Radiation Oncology* **1999**, 9, (1), 48-59.
 14. Xia, P.; Verhey, L. J., Multileaf collimator leaf sequencing algorithm for intensity modulated beams with multiple static segments. *Medical Physics* **1998**, 25, (8), 1424-34.
 15. Shepard, D. M.; Earl, M. A.; Li, X. A.; Naqvi, S.; Yu, C., Direct aperture optimization: a turnkey solution for step-and-shoot IMRT. *Medical Physics* **2002**, 29, (6), 1007-18.
 16. Cotrutz, C.; Xing, L., Segment-Based Dose Optimization Using a Genetic Algorithm. *Physics in Medicine & Biology* **2003**, 48, 2987-2998.
 17. Xing, L.; Li, J. G.; Donaldson, S.; Le, Q. T.; Boyer, A. L., Optimization of importance factors in inverse planning. *Physics in Medicine & Biology* **1999**, 44, (10), 2525-36.
 18. Yang, Y.; Xing, L., Inverse treatment planning with adaptively evolving voxel-dependent penalty scheme. *Med Phys* **2004**, 31, (10), 2839-44.

Appendices -- manuscripts published/submitted

1. Lee L, Le Q, **Xing L**: Retrospective IMRT dose reconstruction based on cone-beam CT and MLC log-file. *International Journal of Radiation Oncology, Biology and Physics*, 70: 634-644, 2008.
2. Chao M, Schreibamnn E, Li T, **Xing L**, Automated contour mapping with a regional deformable model, *International Journal of Radiation Oncology, Biology and Physics*, 70: 599-608, 2008.
3. Mao W, Lee L, **Xing L**, Design of multi-purpose phantom and automated software analysis tool for quality assurance of onboard kV/MV imaging system, *Medical Physics* 35, 1497-1506, 2008.
4. Mao W, Wiersma R, **Xing L**, Fast fiducial detection algorithm for onboard MV and kV imaging systems, *Medical Physics* 35, 1942-1949, 2008.
5. Xie Y, Djajaputra D, King C, Hossain S, Ma L, **Xing L**: Intrafraction motion of prostate in hypofractionated radiation therapy. *International Journal of Radiation Oncology, Biology and Physics* 72, 236-246, 2008.
6. Xie Y, Chao M, Lee P, **Xing L**: Feature-based rectal contour propagation from planning CT to cone beam CT, *Medical Physics* 35, 4450-4459, 2008.
7. Wang J, Li T, Liang Z, **Xing L**, Dose reduction in kV cone beam CT for radiation therapy, *Physics in Medicine and Biology* 53, 2897-2909, 2008.
8. Mao W, Riaz N, Lee L, Wiersma R, **Xing L**, A fiducial detection algorithm for real-time image guided IMRT based on simultaneous MV and kV imaging, *Medical Physics* 35, 3554-3564, 2008.
9. Lee L, Mao W, Xing L, The use of EPID-measured leaf sequence files for IMRT dose reconstruction in adaptive radiation therapy, *Medical Physics* 35, 5019-5029, 2008.
10. Zhu L, Lee L, Ma Y, Ye Y, Mazzeo R, and **Xing L**, Using total-variation regularization for inverse planning with field specific numbers of segments, *Physics in Medicine and Biology* 53, 6653-6672, 2008.
11. Jordan LeNoach, Lei Zhu, Pavel Lougovski, Yunzhi Ma, Lei Xing, IMRT Optimization through Voxel-based Prescription Correction, working paper.
12. Zhu L, and **Xing L**, Search for IMRT solutions with piece-wise constant fluence maps using compressed sensing technique, *Medical Physics*, submitted, 2008.

PHYSICS CONTRIBUTION

RETROSPECTIVE IMRT DOSE RECONSTRUCTION BASED ON
CONE-BEAM CT AND MLC LOG-FILE

LOUIS LEE, PH.D., QUYNH-THU LE, M.D., AND LEI XING, PH.D.

Department of Radiation Oncology, Stanford University School of Medicine, Stanford, CA

Purpose: Head-and-neck (HN) cone-beam computed tomography (CBCT) can be exploited to probe the IMRT dose delivered to a patient taking into account the interfraction anatomic variation and any potential inaccuracy in the IMRT delivery. The aim of this work is to reconstruct the intensity-modulated radiation therapy dose delivered to an HN patient using the CBCT and multileaf collimator (MLC) log-files.

Methods and Materials: A cylindrical CT phantom was used for calibrating the electron density and validating the procedures of the dose reconstruction. Five HN patients were chosen, and for each patient, CBCTs were performed on three separate fractions spaced every 2 weeks starting from the first fraction. The respective MLC log-files were retrieved and converted into fluence maps. The dose was then reconstructed on the corresponding CBCT with the regenerated fluence maps. The reconstructed dose distribution, dosimetric endpoints, and DVHs were compared with that of the treatment plan.

Results: Phantom study showed that HN CBCT can be directly used for dose reconstruction. For most treatment sessions, the CBCT-based dose reconstructions yielded DVHs of the targets close (within 3%) to that of the original treatment plans. However, dosimetric changes (within 10%) due to anatomic variations caused by setup inaccuracy, organ deformation, tumour shrinkage, or weight loss (or a combination of these) were observed for the critical organs.

Conclusions: The methodology we established affords an objective dosimetric basis for the clinical decision on whether a replanning is necessary during the course of treatment and provides a valuable platform for adaptive therapy in future. © 2008 Elsevier Inc.

Cone-beam CT, Head and neck IMRT, MLC log-file, Dose reconstruction, Adaptive radiation therapy.

INTRODUCTION

Two implicit assumptions are made in the current multileaf collimator (MLC)-based intensity modulated radiotherapy (IMRT) process. First, the geometric sizes, shapes, and locations of the targets and organs (internal anatomy) and the geometric topography of the patient are the same as at the time of computed tomography (CT) simulation throughout the treatment course. Second, the delivered fluence maps are the same as the planned ones and delivered by the MLC in an idealized manner. In reality, neither of these assumptions is guaranteed in clinical situations.

Many patients, especially those with head and neck (HN) cancers who undergo fractionated RT course, have marked geometric changes in their internal anatomy and topography during the treatment course, which are attributable to organ deformation, tumor shrinkage, weight loss or a combination

of these (1, 2). These geometric changes might cause undesirable underdosage of the targets and potential overdosage of the critical organs in the vicinity of the targets and lead to a suboptimal treatment outcome. This issue poses a particular concern in HN IMRT cases because of the steep dose gradient that often exists between the boundary of the target and critical organs in an IMRT plan. The use of the three-dimensional (3D) patient model derived from a single planning CT (pCT) set to guide the fractionated RT course is a major hurdle to further advancing the radiation therapy (3). Recently, many in-room imaging modalities including CT-on-rail (4), kilovoltage-CBCT (5, 6), megavoltage cone-beam CT (7), and tomotherapy system (8) have been developed to monitor the geometric changes on a temporal basis. These new imaging modalities are primarily designed to verify and correct the patient's setup in a 3D space with respect to the pCT on the basis of bony landmarks, as well as soft tissue structures,

Reprint requests to: Lei Xing, Ph.D., Department of Radiation Oncology, Stanford University School of Medicine, 875 Blake Wilbur Drive, Room G-204, Stanford, CA 94305-5847. Tel: (650) 498-7896; Fax: (650) 498-4015; E-mail: lei@reyes.stanford.edu

This work was supported by the National Cancer Institute (Grant Nos. 1R01 CA98523 and CA104205), the Komen Breast Cancer Foundation (BCTR0504071), and the Department of Defense (PC040282). Support from the Sir Robert Black Postdoctoral Fel-

lowship and the Hong Kong Hospital Authority Overseas Training Allowance for the first author are also gratefully acknowledged.

Conflict of interest: none.

Acknowledgements—We thank Drs. Sam Brain, Gary Luxton, and Peter Maxim for the Perl program that allowed us to merge the two sets of CBCTs and for useful discussion.

Received June 27, 2007, and in revised form Sept 27, 2007.
Accepted for publication Sept 28, 2007.

before dose delivery. In principle, the CBCT can be further exploited to derive a 3D patient model for dose reconstruction to reflect the dosimetric impact resulting from the residual setup errors and geometric changes occurring over time.

Feasibility studies have been carried out by several research groups to reconstruct the IMRT dose distribution using CBCT and the planned fluence maps from the treatment planning system (TPS) (9–11). This maneuver is based on the second assumption mentioned earlier that the planned fluence maps can be faithfully realized by the delivery system (11). However, in a MLC-based IMRT using the step-and-shoot method, there may be errors associated with the control of leaf motion and fractional monitor unit (MU) delivery such as overshoot, undershoot segmental MU, dropped segments, and beam delivery during leaf motion (12–15). These factors might affect the dose delivery and result in a delivered fluence map different from the planned one. A more pragmatic approach in reconstructing the delivered dose is to use the fluence maps actually delivered for the treatment; therefore, we propose retrieving the MLC log-file that records the leaf position of each individual leaf and the fractional MU status during the delivery of IMRT from the MLC workstation and deriving the delivered fluence map from it.

The objective of this study is twofold: (1) to establish a methodology and procedures to reconstruct the dose delivered to a patient on a series of kV-CBCTs taken during a treatment course using the delivered fluence maps derived from the corresponding MLC log-files and (2) to study the potential dosimetric impact on the intended treatment plan taking into account the patient's geometric changes over time, residual setup errors, and the inherent delivery errors associated with the MLC. This work can serve as a platform for implementing a workflow in reconstructing the IMRT delivered dose and providing the necessary dosimetric information needed to modify the treatment plan, if indicated, on the basis of the accumulated dose given to the patient.

METHODS AND MATERIALS

CBCT image acquisition

The CBCT images in this study were acquired by the onboard imager (OBI) integrated in a Trilogy medical linear accelerator (Varian Medical Systems, Palo Alto, CA). The OBI system is mounted on the gantry of the linear accelerator perpendicular to the beam axis of the MV beam by robotic arms. The OBI system consists of a kV X-ray tube assembly at one end and an amorphous silicon flat-panel image receptor (39.7 cm × 29.8 cm) (Varian 4030CB flat panel) facing the X-ray tube at the other end in the full-extended configuration. The focus of the X-ray tube and the center of the image receptor are at 100 cm and 50 cm from the isocenter, respectively, resulting in a source-to-imager distance of 150 cm. The OBI isocenter coincides with the MV treatment isocenter within ± 1.5 mm and is routinely checked in the weekly quality assurance procedure (16). The CBCT can be acquired in two modes, the “half-fan” mode and “full-fan” mode. The half-fan mode is designed to increase the field-of-view (FOV) beyond 24 cm and was used exclusively in this study. In the half-fan mode, the image receptor is displaced laterally by 14.8 cm, and the blades of the X-ray tube are offset to cover the detector area. A half-bowtie filter is used in this

mode. A half-fan projection image is acquired for each acquirement angle for the 364° gantry rotation in about 60 s, resulting in a total of 640–700 projections (16). Only part of the object is viewed in one half-fan projection; the other part of the object is viewed in the half-fan projection from the opposite position. The entire object is reconstructed by using the projections acquired 180° apart. The FOV and longitudinal coverage in this mode are 45 cm in diameter and 15 cm in length, respectively. A total of 640 projections are acquired and reconstructed into CBCT images of 512 × 512 matrix.

Electron density calibration

Electron density calibration was performed by scanning a phantom with inserts of known relative electron densities with respect to water (ρ_e^w) and calibrating the ρ_e^w against the Hounsfield units (HU) of the inserts. A 20-cm cylindrical CT phantom, Catphan-600 with CTP404 module (Phantom Laboratory, NY), was used for the electron density calibration for the pCT (GE Discovery-ST PET/CT scanner, Milwaukee, WI) and CBCT. This phantom was chosen because it has a size and shape comparable to a patient's HN region. Designated scanning parameters for HN patients were chosen for the pCT and similar scanning parameters were applied as far as possible for the CBCT; the slice thickness used was 2.5 mm. The half-fan CBCT mode (half-bowtie mounted) was used to give a FOV of 45 cm, comparable to that of 45–50 cm used in pCT to include the lower neck region. Both CT sets had a pixel size of about 1 mm in a 512 × 512 matrix image. The techniques for the CBCT used were 125 kVp, 80 mA, and 25-ms pulse, which are precalibrated for clinical use. The HUs of the inserts for both the pCT and CBCT were measured from the acquired images and plotted against the known ρ_e^w . The vertical and horizontal HU profiles of the same images were also plotted and compared. The temporal CBCT HU stability has been investigated by our group, and there is no significant fluctuation observed over a period of 8 weeks (10).

MLC log-file retrieval and processing

The Trilogy is equipped with the Millennium 120-leaf MLC (Varian Medical Systems) capable of IMRT delivery. During a step-and-shoot IMRT delivery, the MLC workstation logs the position of each individual leaf and the fractional MUs delivered every 50 ms; it also produces two separate log-files for each leaf bank (A and B sides). A split IMRT field results in four such MLC log-files. These automatically generated MLC log-files have been validated to represent accurately the actual IMRT delivery (17, 18). For this study, the MLC log-files were retrieved from the MLC workstation after the treatment session in which the CBCT was performed; the leaf positions and delivered fractional MUs were extracted and converted to leaf sequence files by software developed in-house, written in Visual Basic 6.0 code (Microsoft, Bellevue, WA). The leaf sequence files were then imported back into the TPS to regenerate the delivered fluence maps, which were used in the dose reconstruction on the CBCT-derived 3D patient model (Fig. 1). To access the practicability of the retrieval and conversion procedures, the MLC log-files of a HN IMRT patient were retrieved, and the subsequently regenerated fluence maps were attached to the original treatment fields in the IMRT plan. The dose was reconstructed using the original set of pCT; the resultant dose distribution and dose volume histograms (DVH) of the target and organs were then compared with those from the original plan to see the discrepancy due solely to the difference in the fluence maps used.

Phantom study

Planning CT and CBCT were acquired for the Catphan-600 phantom using the techniques previously described. Both sets of images

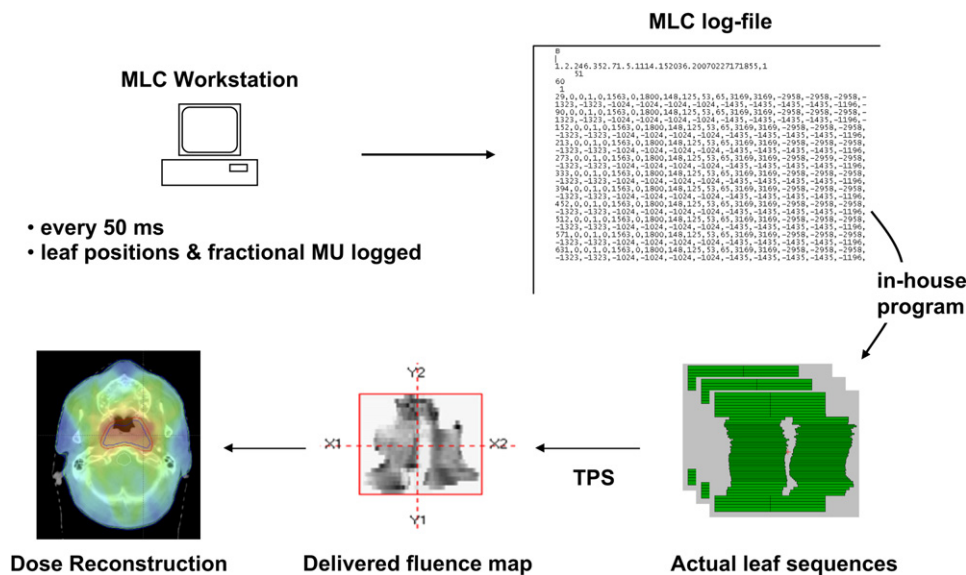


Fig. 1. The workflow from the retrieval of the multileaf collimator (MLC) log-file to the regeneration of the delivered fluence map and dose reconstruction. MU = monitor unit; TPS = treatment planning system.

were imported into the Eclipse TPS (Varian Medical Systems) for feasibility study. A hypothetical identical target in a shape of cuboid ($4 \times 4 \times 2$ cm) was created in the center of the phantom for the two image sets. An isocentric plan with two orthogonal 6-MV fields (Varian scale 180° and 90° gantry angle) each measuring 10×10 cm was applied to the center of the phantom in each image set. The Anisotropic Analytical Algorithm (AAA) implemented in the Eclipse TPS was adopted for dose calculation. The electron calibration curve of the pCT, which had been commissioned for the use of dose calculation in TPS, was applied to both image sets for dose calculation; this is based on the findings from the electron density calibration scans performed on the pCT and CBCT earlier. A five-field IMRT plan from an HN patient was also applied to two image sets of the phantom to validate further the use of CBCT for dose calculation. The resultant dose distributions, orthogonal dose profiles, and DVHs of the hypothetical target from the pCT- and CBCT-based calculations were compared.

Patient CBCT scheduling and acquisition

Five HN IMRT patients were chosen for this work. The patient was setup under the guidance of kV orthogonal planar imaging routinely using the OBI. The CBCT was then acquired using the techniques previously described. There was no attempt to use the CBCT for 3D-3D setup correction because 3D-3D setup verification has not yet been adopted as a routine clinical procedure at our clinic. Because of the limited coverage of the CBCT in the longitudinal direction, two sets of CBCT 10 cm apart in this direction were taken for each patient. The two CBCT image sets were merged together

before they were imported into the TPS. Three CBCTs at 2-week intervals starting from the first or second fraction in a 30-fraction treatment course (Fig. 2) were taken for the patients. There is usually a 2-week gap from the pCT to the first treatment; the first CBCT not only serves as a starting time point to see the temporal geometric changes but also shows the geometric changes that might have occurred during the 2-week gap before the treatment commences. The number of CBCTs taken and the time interval between them are a compromise based on the consideration of imparted CBCT dose, workload, and the progress of anatomic changes that manifest on the CBCT. The acquired CBCTs and the retrospective dose reconstructions were used only for research purposes and not for altering the original treatment or clinical practice.

Procedures for dose reconstruction

After each treatment, the MLC log-files and the CBCT image sets were retrieved. The merged HN CBCT was imported into the Eclipse TPS; a 3D patient model was generated from the CBCT and fused to the pCT 3D patient model using a rigid-body image registration technique based on mutual information of the whole CT volumes. This image registration was used only for mapping the targets and organ contours from the pCT image set onto the CBCT image set. These initial mapped contours were edited by the same physician who did the contouring for the original plan, if necessary, to adapt to changes in the patient's anatomy due to tumour shrinkage, organ deformation, or weight loss. After the CBCT image set had been finalized with the corrected contours, it was decoupled from the registration to retain its own geometric status, including

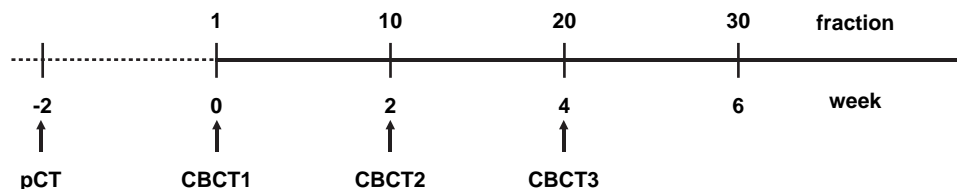


Fig. 2. Timeline showing the scheduling of the cone-beam computed tomography (CBCT) and the time of planning computed tomography (pCT).

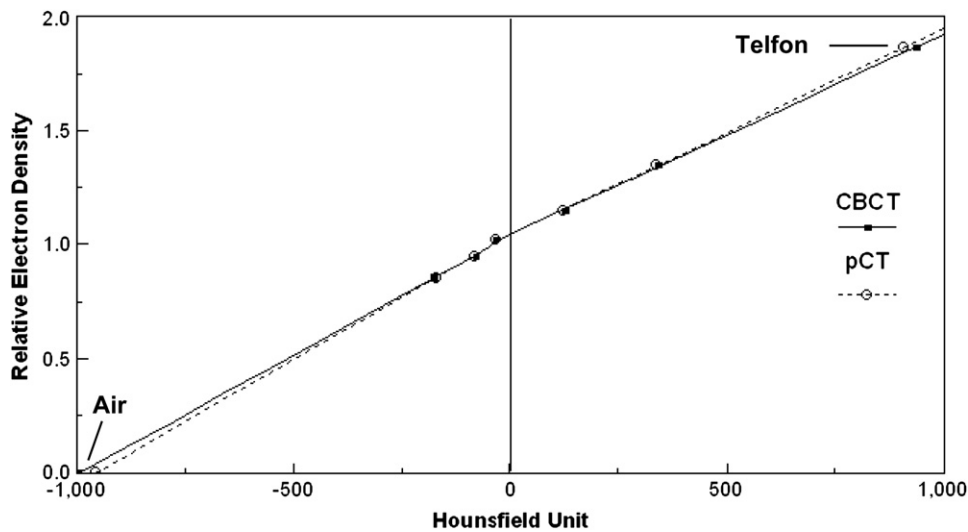


Fig. 3. The electron density calibration curves for the planning computed tomography (pCT) and cone-beam computed tomography (CBCT).

the residual setup errors for the subsequent dose reconstruction. The configuration including the beam angulations and field sizes of the original treatment plan were copied to the CBCT and placed at the same isocenter of the pCT. Because the MLC log-files are logged for each independent delivery, a split IMRT field would have to be considered as two independent subfields, each with its own derived fluence map when copied to the CBCT plan. A typical HN

IMRT plan with seven split-fields would result in 14 individual subfields each with its corresponding derived fluence maps. The delivered MUs for each subfield were entered when the fluence map was regenerated. The dose reconstruction was then calculated on the CBCT-derived patient model. The resultant dose distributions, dosimetric endpoints, and DVHs of the target and critical organs were compared with those of the original plan and the other two CBCTs.

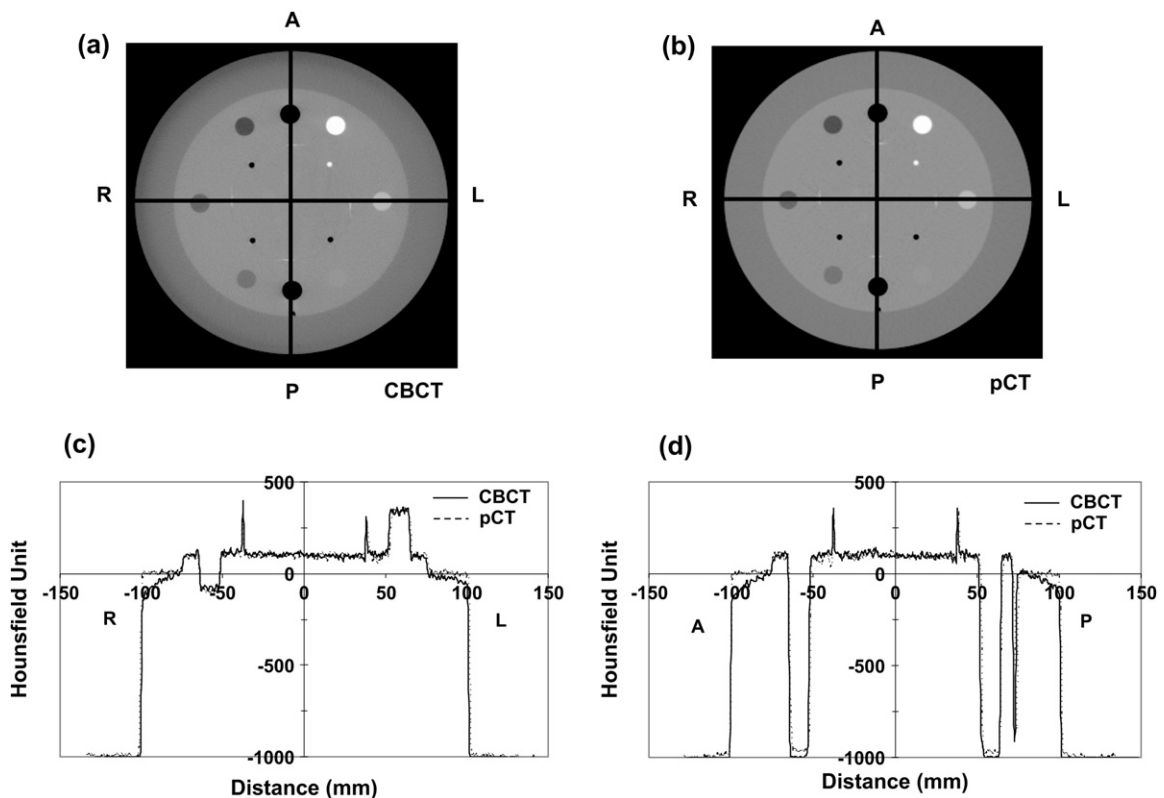


Fig. 4. An axial slice through the inserts from (a) the cone-beam computed tomography (CBCT) and (b) the planning computed tomography (pCT). (c) The horizontal Hounsfield unit (HU) profiles along the lines right-to-left (RL) for the CBCT and pCT. (d) The vertical HU profiles along the lines anteroposterior (AP) for the CBCT and pCT.

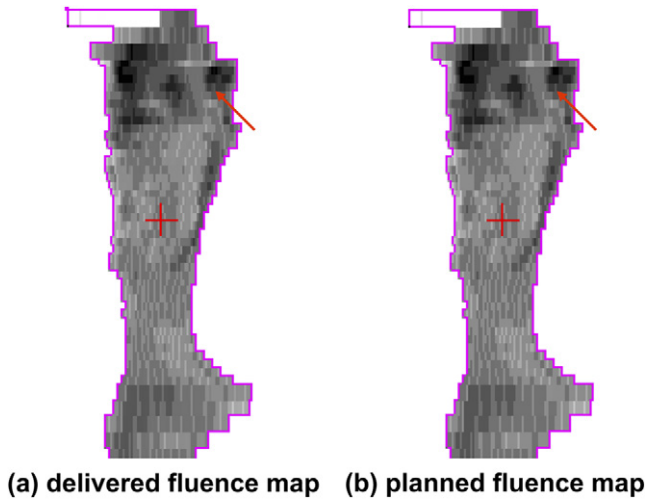


Fig. 5. No discernible difference was seen between the two fluence maps except at the region indicated by the arrow.

RESULTS

Electron density calibration

Figure 3 shows the relative electron density calibration curves for the pCT and CBCT. No significant difference in the calibrations was noted over the range of HUs. The differences in the HUs of the inserts for the two modes of CT were less than 10 HUs, except for the air and Teflon inserts; however, the maximum relative difference was only 5% and 3% for the air and Teflon, respectively.

Figures 4a and 4b show the axial images of the phantom through the inserts from the CBCT and pCT, respectively.

The horizontal and vertical HU profiles along the lines right-to-left (RL) and anteroposterior (AP) for the CBCT and pCT are compared in Figs. 4c and 4d. The CBCT and pCT HU profiles were in agreement to within 10% except that the CBCT showed a larger fluctuation of HU and reduced HUs at the periphery of the phantom.

Derivation of fluence maps from the MLC log-files

Figures 5a and 5b show a delivered fluence map derived from the MLC log-files and the corresponding planned fluence map from the TPS, respectively. There was essentially no discernible difference in the intensity levels between the two maps except at the region indicated by the arrow. Figures 6a and 6b depict the dose distribution on an axial slice calculated on the same pCT set using the delivered and planned fluence maps; their corresponding DVHs of the targets and critical organs are shown in Figs. 6c and 6d. There was essentially no discernible difference in the dose distributions on the axial slice. The DVHs of the planning tumour volumes (PTV), gross tumour volumes (GTV), and critical organs of the plan based on delivered fluence maps overlapped with the ones from the plan based on planned fluence maps.

Phantom study

For the open beam dose calculation, the dose distributions on the same axial slice from the CBCT- and pCT-based dose calculations were essentially identical; the difference in dose maximum was less than 0.5% (Figs. 7a and 7b). The horizontal and vertical dose profiles for the CBCT- and pCT-based calculations were in agreement to within 1% except at the field

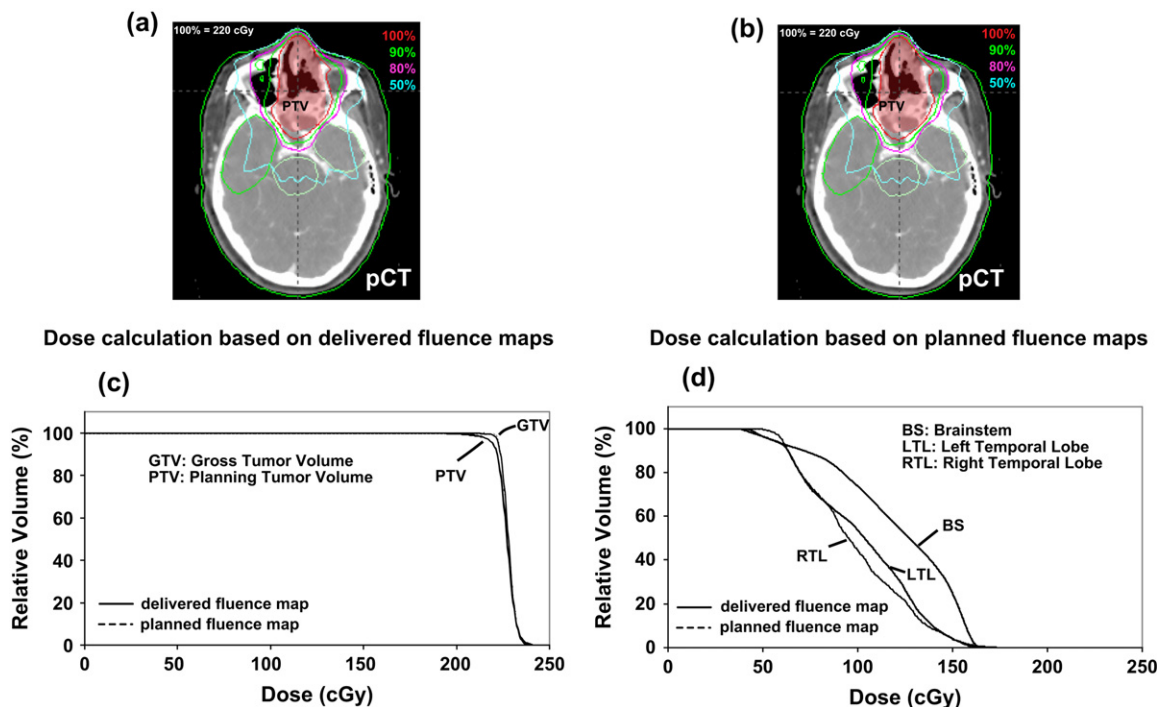


Fig. 6. Intensity-modulated radiation therapy dose distribution on an axial slice from the planning computed tomography for the dose calculations based on (a) the delivered and (b) the planned fluence maps. Dose volume histograms of the gross tumor volume (GTV), planning tumor volume (PTV) (c), and the critical organs (d) for the two calculations.

edges, where there was a maximum difference of 3%. (Fig. 7c). The DVHs of the hypothetical target from the two calculations essentially overlapped (Fig. 7d). The same comparisons were performed for the IMRT dose calculation. There was also no significant difference in the dose distributions on the axial slices shown and in the dose maximum points (Figs. 8a and 8b). Both the dose profiles and the DVHs from the two calculations were in agreement to within 2% (Figs. 8c and 8d).

Patient study

Patients with PTV in the head region. Of the three patients (A, B, and C) in the study, two had nasopharyngeal carcinoma (NPC), the other one had esthesioneuroblastoma surgically debulked, with residual tumour at the base of skull to be treated. The PTVs of these patients received 6,600 cGy in 30 fractions at 220 cGy per fraction. For the patients studied here, we observed that the change in the shape of the targets had already occurred during the 2-week waiting period. These were then followed by a progressive change in the shape and size of the targets during the treatment. The dosimetric comparisons of the pCT- with CBCT-based plans per fraction for the three patients are shown in Table 1. For the PTV, the dose to 95% of the volume (D_{95}) and the volume receiving at least 93% of the prescribed dose (V_{93}) of the pCT were in agreement to within 3% with those of the CBCTs. For the GTV, the dose to 99% of the volume (D_{99}) also showed the same agreement between the pCT and CBCTs. The differences in the dosimetric endpoints of the critical organs for the pCT and CBCTs were generally in

the range of 10%; however, the difference could be as high as 15%–20% in some individual treatment sessions for the parotid gland, optic track, and the temporal lobe when they were in extreme proximity with the PTV. The comparisons of the dose distributions and the DVHs between the pCT- and CBCT-based plans for one of the patients are shown in Fig. 9 for illustrative purposes.

Patients with PTV in the neck region. The two other patients (D and E) we selected had primary tumors in the hypopharynx and the lacrimal gland; both had undergone surgery. There were two PTVs defined for these patients: PTV1 for the tumor bed and PTV2 for the involved neck. PTV1 and PTV2 received 6,000 cGy and 5,400 cGy, respectively, in 30 fractions at 200 cGy per fraction. Because of the limited coverage of the CBCT, the PTV1 of one of the patients (E) was not completely included and was not compared. The main emphasis here was to study the dosimetric impact of the change of the neck contour on the PTV2. From the first CBCT, the neck contours of these patients had a marked change from the pCT, and the change was slight throughout the progression of the treatment up to the third CBCT (Fig. 10). For the PTV1 of the patient D, D_{95} and V_{93} of the pCT-based plan agreed well with those of the CBCT-based plans (Table 2). For the PTV2 of both patients, D_{95} and V_{93} showed agreement to within 3% between the pCT- and CBCT-based plans. The differences in the dosimetric endpoints of the critical organs for the pCT- and CBCT-based plans were generally in the range of 5% (Table 2). The comparisons of the dose distributions and the DVHs between the pCT- and CBCT-based plans for one of these patients are shown in Fig. 10.

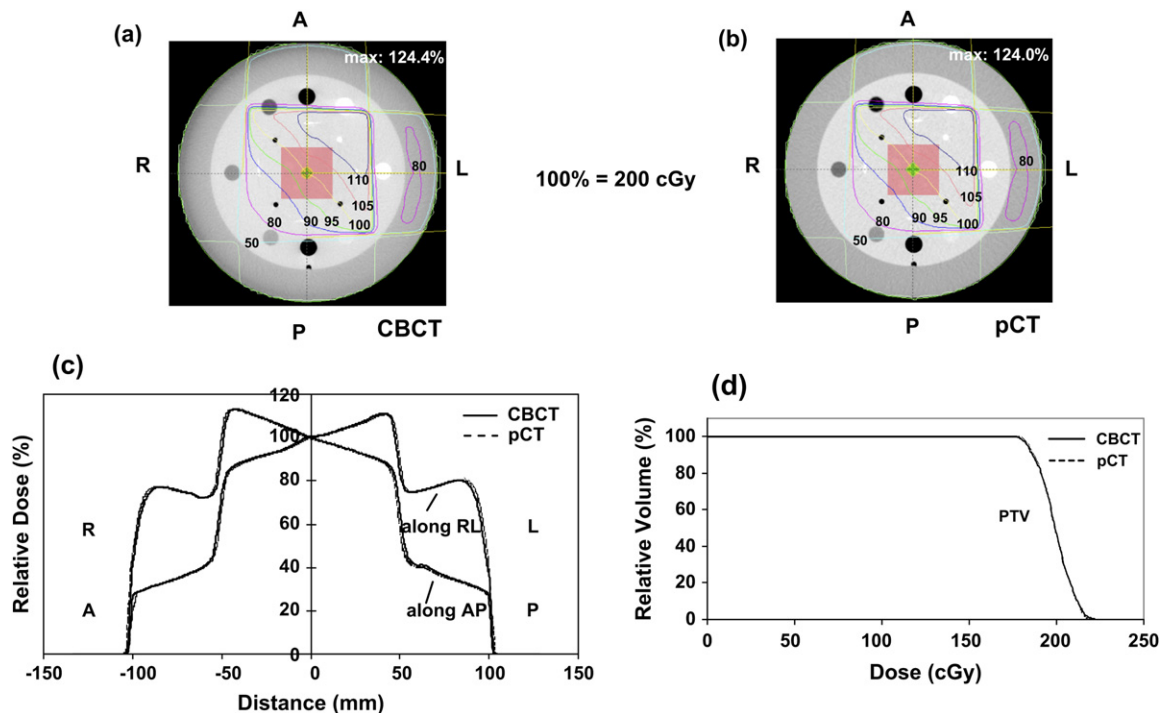


Fig. 7. Dose distribution from two orthogonal 10×10 cm open fields on the axial slice with (a) the cone-beam computed tomography (CBCT)-based and (b) the planning computed tomography (pCT)-based dose calculations. (c) Dose profiles along lines right-to-left (RL) and anteroposterior (AP) for CBCT- and pCT-based calculations. (d) Dose volume histograms of the hypothetical target for the CBCT-based and pCT-based calculations.

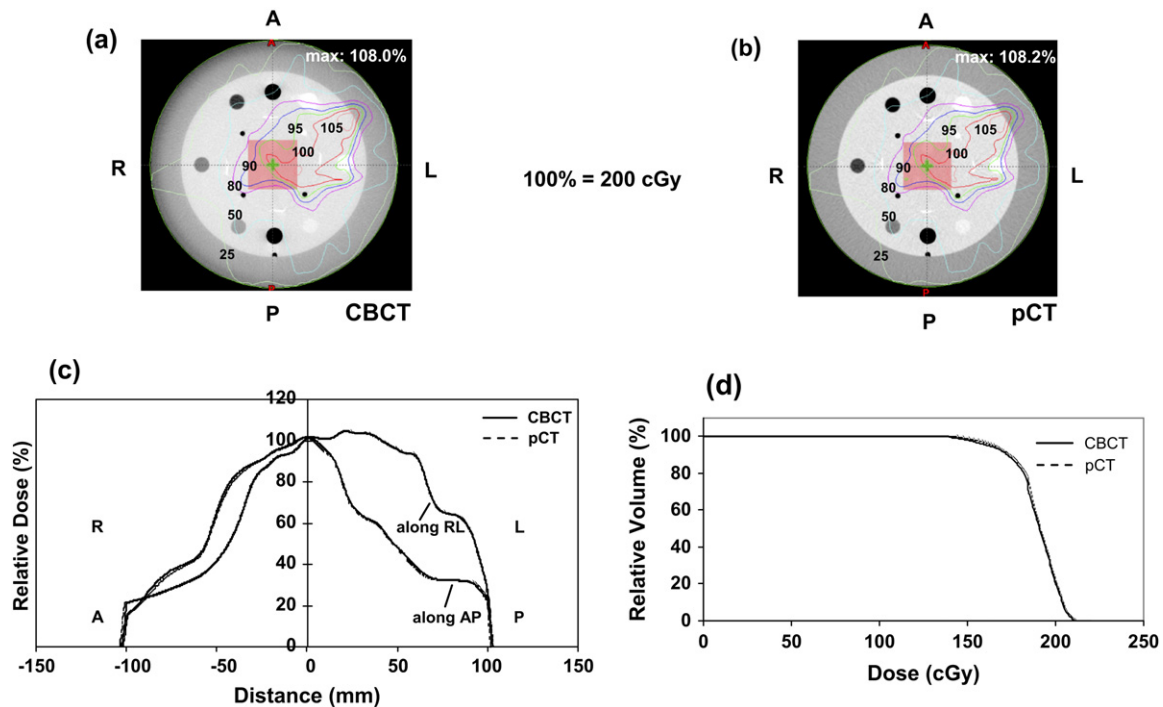


Fig. 8. Intensity modulated radiation therapy (IMRT) dose distribution on the axial slice with (a) the cone-beam computed tomography (CBCT)-based and (b) the planning computed tomography (pCT)-based dose calculations. (c) Dose profiles along lines right-to-left (RL) and anteroposterior (AP) for CBCT-based and pCT-based calculations. (d) Dose volume histograms of the hypothetical target for the CBCT-based and pCT-based calculations.

DISCUSSION

The advent of onboard CBCT expands our horizon in correcting the daily setup errors and probing the geometric change of a patient's anatomy in his or her treatment position during the course of radiation treatment. We have shown that

the CBCT could be further exploited for dose reconstruction to provide the dosimetric information necessary for replanning or reoptimization should there be significant deviations in the delivered dose from the intended plan. The main hurdle in preventing us from using the CBCT for dose reconstruction

Table 1. Comparison of dosimetric endpoints of the targets and critical organs between the pCT- and the serial CBCT-based plans for the patients A, B, and C (Prescription: 6600 cGy to PTV in 30 fractions at 220 cGy per fraction; values shown below are for one fraction)

Dosimetric end point	Patient A				Patient B				Patient C			
	pCT	CBCT1 pCT	CBCT2 pCT	CBCT3 pCT	pCT	CBCT1 pCT	CBCT2 pCT	CBCT3 pCT	pCT	CBCT1 pCT	CBCT2 pCT	CBCT3 pCT
PTV, D ₉₅ (cGy)	220	0.98	0.99	0.98	220	0.97	0.97	0.99	220	1.01	1.00	1.01
PTV, V ₉₃ (%)	99	0.99	1.00	0.99	100	0.99	0.98	1.00	100	1.00	1.00	1.00
GTV, D ₉₉ (cGy)	220	0.97	0.99	0.98	219	1.00	1.00	0.99	219	1.01	1.00	1.01
BS, D _{max} (cGy)	165	1.06	1.04	1.04	156	1.02	1.09	1.03	153	1.08	1.02	1.03
SC, D _{max} (cGy)	—	—	—	—	139	0.97	0.97	0.99	150	1.01	1.00	1.01
OC, D _{max} (cGy)	164	1.01	1.13	1.06	—	—	—	—	156	1.08	0.97	1.01
R ON, D _{max} (cGy)	164	0.93	0.97	0.92	—	—	—	—	152	1.04	0.93	1.00
L ON, D _{max} (cGy)	170	1.08	0.84	1.10	—	—	—	—	156	1.01	0.98	0.98
R PARO, D _{mean} (cGy)	—	—	—	—	113	1.05	0.85	1.07	90	0.97	0.99	1.19
L PARO, D _{mean} (cGy)	—	—	—	—	141	1.07	1.12	1.03	108	0.97	1.08	1.03
R TL, D _{max} (cGy)	175	1.00	1.00	0.98	160	0.95	0.92	0.97	—	—	—	—
L TL, D _{max} (cGy)	175	1.09	1.19	1.23	158	1.03	1.06	0.98	—	—	—	—

Abbreviations: BS = brainstem; CBCT = cone-beam computed tomography; D₉₅ = dose to 95% of the volume; D₉₉ = dose to 99% of the volume; D_{max} = maximum dose; D_{mean} = mean dose; GTV = gross target volume; L = left; OC = optic chiasm; ON = optic nerve; PARO = parotid gland; pCT = planning computed tomography; PTV = planning target volume; R = right; SC = spinal cord; TL = temporal lobe; V₉₃ = volume receiving at least 93% of the prescribed dose.

Note: For each patient, the first column lists the value for the pCT-based plan, and the second to fourth columns show the ratio of the values of the subsequent CBCT- to pCT-based plan for easier comparison. Blank cells indicate that the organ is not included in the CBCT coverage because of PTV site.

is that the CBCT does not always warrant an image quality comparable to conventional CT because of the scatter environment encountered in CBCT (19). Another issue is the provision of an appropriate electron density calibration; large variations of HU were observed for the same material under different scanning conditions and parameters used in CBCT (9, 10). We mitigated these problems by restricting the electron density calibration to designated scanning parameters and using a cylindrical phantom with a caliber close to a patient's HN region. By scanning the Catphan in the similar scanning parameters in pCT and in CBCT, we found that the electron density calibration curves had no significant difference except for that of the air and Teflon; however, these 3%–5% HU difference will not transform into clinically significant difference in doses (<1%) (20, 21). This explained why the fluctuating HUs and peripheral reduced HUs seen in the CBCT phantom images (Figs. 4c and 4d) did not lead to any significant difference in the dose profiles (Figs. 7c and 8c) when we validated the CBCT for dose calculation. Therefore, we adopted to use the commissioned pCT electron density calibration for the CBCT dose calculation; this has also been proposed by Yoo and Yin on the basis of the same argument (9). The good agreement of the DVHs of the hypothetical targets for the pCT- and CBCT-based dose calculations (Figs. 7d and 8d) further justifies our method. However, the ultimate solution is the continual improvement in the image reconstruction algorithm for the CBCT making it more

consistent and reliable under different scanning conditions and parameters (19). The results from the phantom study strongly supported that the CBCT obtained under the proposed scanning condition could be directly used for dose reconstruction.

The MLC log-files have been validated to faithfully reflect the actual delivery process of a MLC-based IMRT (17, 18); this serves to be a powerful tool in verifying or reconstructing the IMRT dose delivered to the patients. The use of the MLC sequence files from the original plan instead of the MLC log-files in dose reconstruction, which most other research groups did (9, 11, 22, 23), rests on the assumption the treatment was indeed delivered as intended. Our approach of using MLC log-file is more realistic and has the potential to take MLC delivery errors into account. This approach is straightforward and easy to implement in a clinic.

From the serial CBCTs performed on the patients, we noticed that the first CBCT provided valuable geometric information about the target and the patient's topography at the start of the treatment. The other two CBCTs monitored the geometric change of the internal anatomy and the dose reconstruction of these provided us the necessary dosimetric information necessary for the decision of replanning or reoptimizing. From the patient study, we found that there was no significant compromise (<3%) on the target coverage even though there was evidence of geometric changes during the course of treatment. The impact of these geometric changes,

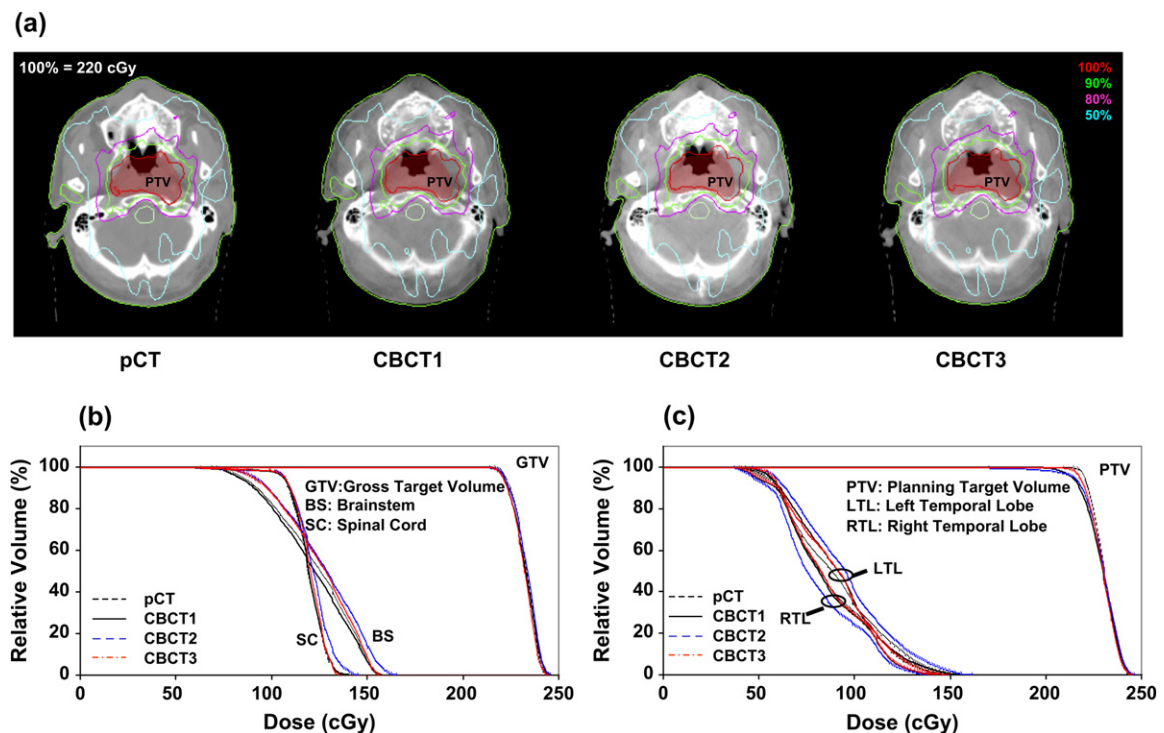


Fig. 9. (a) Geometric change of the target of patient B was noted in the first cone-beam computed tomography (CBCT). The CBCT dose reconstructions yielded dose distributions close to that of the original plan. (b) The dose volume histograms (DVHs) of the gross tumor volume (GTV) for the planning computed tomography (pCT)-based and cone-beam computed tomography (CBCT)-based plans essentially overlapped. (c) The coverage of the planning tumor volume (PTVs) in the CBCTs were slightly compromised when compared with that of the pCT. The dose volume histograms of the critical organs of the CBCT-based plans deviated from those of the original plan in various degrees (b) and (c).

however, largely affected the dose deposited to the critical organs in the vicinity of the target. This might be due to the steep dose gradient that commonly exists between the target and the critical organs in an IMRT plan, and any deviated geometric change, whether due to the residual setup error, tumor shrinkage, organ deformation or patient weight loss, would result in a considerable change (up to 20%) in the dose received by these critical organs. Another less subtle reason is the inaccuracy in contouring these organs in the CBCT that was not supplemented by the contrast enhancement or magnetic resonance imaging (MRI) fusion as the pCT had been. This is particularly crucial when one comes to an organ of small volume such as the optic nerve or optic chiasm, for which a slight error in contouring would cause a drastic change of the dose received (Table 1). In the patients with targets in the neck region, despite the fact that there were marked changes of the neck contour on the CBCTs compared with the pCT, the coverage of the neck target (PTV2) was not significantly compromised, probably because these patients' original PTV2 had been markedly modified to be confined to the neck and still received adequate dose coverage. Because of the small number of patients in this work, it is not recommended that these findings be generalized to the dosimetric impact of geometric changes. Nevertheless, this work gives us some insight how important and useful the on-treatment serial CBCTs could be in adapting the radiation treatment in future.

Although we have set up a methodology and procedures to reconstruct the IMRT dose on CBCT, we still face some practical issues in implementing it into clinical practice. First, we must determine how many CBCTs should be taken, and how often, so that a clinically meaningful intervention can rely on. The contouring of the targets and organs in all the CBCTs is another practical problem we encounter; until we have a reliable deformable image registration that can directly propagate the contours from the pCT to CBCTs (24, 25), it is difficult to handle contouring in a large number of CBCT sets. Furthermore, it has always been technically challenging to interpret dose distributions and DVHs and infer from them the cumulative doses from various sets of 3D patient models with changing internal anatomy (26, 27). The DVH definition may need to be redefined to handle the changing target (28). We must identify predictive indices on the basis of all the dosimetric information gained from the CBCTs to guide us in making a clinical decision regarding when and how a treatment can be adapted appropriately and efficaciously should the need arise. All these factors need to be addressed in future research. An ideal paradigm that integrates a seamless workflow from simulation to planning, verification, offline or on-line replanning (15), and finally delivery is essential for the success of adaptive radiation therapy (26). This study has demonstrated the fundamental steps required for the full implementation of using CBCT for this ideal paradigm.

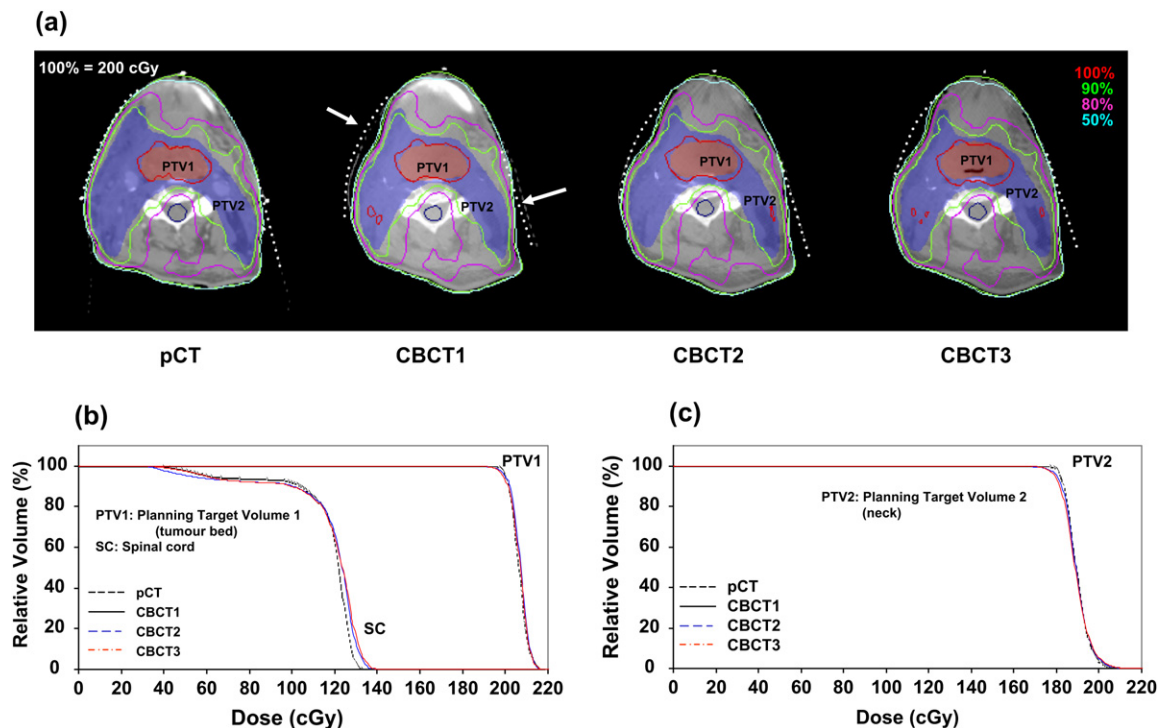


Fig. 10. (a) Marked changes of the neck contour (white arrows) and planning tumor volume 2 (PTV2) were noted in the first cone-beam computed tomography (CBCT) of patient D. The CBCT dose reconstructions otherwise yielded dose distributions close to that of the original plan when geometric changes were taken into account. (b) The dose volume histograms (DVHs) of the PTV1 for the planning computed tomography (pCT)- and CBCT-based plans essentially overlapped. (c) The coverage of the PTV2s in the CBCT-based plans were slightly compromised compared with that of the pCT-based plan. The DVHs of the spinal cord of the CBCT-based plans deviated from those of the original plan (b).

Table 2. Comparison of dosimetric endpoints of the targets and critical organs between the pCT- and the serial CBCT-based plans for patients D and E (prescription: 6,000 cGy to PTV1 and 5,400 cGy to PTV2 in 30 fractions at 200 cGy per fraction; values shown are for one fraction)

Dosimetric end point	Patient D				Patient E			
	pCT	$\frac{CBCT1}{pCT}$	$\frac{CBCT2}{pCT}$	$\frac{CBCT3}{pCT}$	pCT	$\frac{CBCT1}{pCT}$	$\frac{CBCT2}{pCT}$	$\frac{CBCT3}{pCT}$
PTV1, D ₉₅ (cGy)	200	1.00	1.00	1.00	—	—	—	—
PTV1, V ₉₃ (%)	100	1.00	1.00	1.00	—	—	—	—
PTV2, D ₉₅ (cGy)	182	0.99	0.99	0.99	179	0.98	0.98	0.97
PTV2, V ₉₃ (%)	100	1.00	1.00	1.00	100	0.99	1.00	0.99
BS, D _{max} (cGy)	—	—	—	—	124	0.96	0.97	0.92
SC, D _{max} (cGy)	135	1.05	1.04	1.05	109	1.05	1.04	1.00

Abbreviations: BS = brainstem; CBCT = cone-beam computed tomography; D₉₅ = dose to 95% of the volume; D_{max} = maximum dose; pCT = planning computed tomography; PTV1 = planning target volume 1 (tumour bed); PTV2 = planning target volume 2 (neck); SC = spinal cord; V₉₃ = volume receiving at least 93% of the prescribed dose.

Note: For each patient, the first column lists the value for the pCT-based plan, and the second to fourth columns show the ratio of the values of the subsequent CBCT- to pCT-based plans for easier comparison. Blank cells indicate that the target or organ is not included in the CBCT coverage.

CONCLUSIONS

In this study, we established methodology and procedures to reconstruct the IMRT dose delivered on the basis of a series of on-treatment CBCTs and the MLC log-files, pragmatically taking into account geometric changes of patients' anatomy over time and the potential inaccuracy in the IMRT delivery.

Dose reconstruction is valuable for examining the actual dose delivered to a patient at a particular fraction. This maneuver affords an objective dosimetric basis for clinical decision making on whether replanning or reoptimization is necessary during the course of treatment. It also provides a valuable platform for adaptive radiation therapy in future.

REFERENCES

- Barker JL Jr., Garden AS, Ang KK, *et al.* Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. *Int J Radiat Oncol Biol Phys* 2004;59:960–970.
- Sobel S, Rubin P, Keller B, *et al.* Tumor persistence as a predictor of outcome after radiation therapy of head and neck cancers. *Int J Radiat Oncol Biol Phys* 1976;1:873–880.
- Yan D, Lockman D, Martinez A, *et al.* Computed tomography guided management of interfractional patient variation. *Semin Radiat Oncol* 2005;15:168–179.
- Ma CM, Paskalev K. In-room CT techniques for image-guided radiation therapy. *Med Dosim* 2006;31:30–39.
- Oldham M, Letourneau D, Watt L, *et al.* Cone-beam-CT guided radiation therapy: A model for on-line application. *Radiother Oncol* 2005;75:271–278.
- Jaffray DA, Siewerdsen JH, Wong JW, *et al.* Flat-panel cone-beam computed tomography for image-guided radiation therapy. *Int J Radiat Oncol Biol Phys* 2002;53:1337–1349.
- Pouliot J, Bani-Hashemi A, Chen J, *et al.* Low-dose megavoltage cone-beam CT for radiation therapy. *Int J Radiat Oncol Biol Phys* 2005;61:552–560.
- Mackie TR, Kapatoes J, Ruchala K, *et al.* Image guidance for precise conformal radiotherapy. *Int J Radiat Oncol Biol Phys* 2003;56:89–105.
- Yoo S, Yin FF. Dosimetric feasibility of cone-beam CT-based treatment planning compared to CT-based treatment planning. *Int J Radiat Oncol Biol Phys* 2006;66:1553–1561.
- Yang Y, Schreiber E, Li T, *et al.* Evaluation of on-board kV cone beam CT (CBCT)-based dose calculation. *Phys Med Biol* 2007;52:685–705.
- Langen KM, Meeks SL, Poole DO, *et al.* The use of megavoltage CT (MVCT) images for dose recomputations. *Phys Med Biol* 2005;50:4259–4276.
- Ezzell GA, Chungbin S. The overshoot phenomenon in step-and-shoot IMRT delivery. *J Appl Clin Med Phys* 2001;2:138–148.
- Xia P, Chuang CF, Verhey LJ. Communication and sampling rate limitations in IMRT delivery with a dynamic multileaf collimator system. *Med Phys* 2002;29:412–423.
- Wiersma R, Xing L. Examination of geometric and dosimetric accuracies of gated step-and-shoot IMRT. *Med Phys* 2007;34:3962–3970.
- Litzenberg DW, Hadley SW, Tyagi N, *et al.* Synchronized dynamic dose reconstruction. *Med Phys* 2007;34:91–102.
- Yoo S, Kim GY, Hammoud R, *et al.* A quality assurance program for the on-board imagers. *Med Phys* 2006;33:4431–4447.
- Li JG, Dempsey JF, Ding L, *et al.* Validation of dynamic MLC-controller log files using a two-dimensional diode array. *Med Phys* 2003;30:799–805.
- Stell AM, Li JG, Zeidan OA, *et al.* An extensive log-file analysis of step-and-shoot intensity modulated radiation therapy segment delivery errors. *Med Phys* 2004;31:1593–1602.
- Zhu L, Starlack J, Bennett NR, Li T, Xing L, Fahrig R. Improved scatter correction for x-ray conebeam CT using primary modulation. In: Hsieh J, Flynn M, editors. *Proceedings of SPIE: Medical Imaging 2007: Physics of Medical Imaging*, San Diego, CA; 2007.
- Cozzi L, Fogliata A, Buffa F, *et al.* Dosimetric impact of computed tomography calibration on a commercial treatment planning system for external radiation therapy. *Radiother Oncol* 1998;48:335–338.

21. Guan H, Yin FF, Kim JH. Accuracy of inhomogeneity correction in photon radiotherapy from CT scans with different settings. *Phys Med Biol* 2002;47:N223–N231.
22. Hansen EK, Bucci MK, Quivey JM, *et al.* Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2006;64:355–362.
23. Morin O, Chen J, Aubin M, *et al.* Dose calculation using megavoltage cone-beam CT. *Int J Radiat Oncol Biol Phys* 2007;67:1201–1210.
24. Xing L, Thorndyke B, Schreibmann E, *et al.* Overview of image-guided radiation therapy. *Med Dosim* 2006;31:91–112.
25. Schreibmann E, Chen GT, Xing L. Image interpolation in 4D CT using a BSpline deformable registration model. *Int J Radiat Oncol Biol Phys* 2006;64:1537–1550.
26. Zerda Adl, Armbruster B, Xing L. Formulating adaptive radiation therapy (ART) treatment planning into a closed-loop control framework. *Phys Med Biol* 2007;52:4137–4153.
27. Yan D, Jaffray DA, Wong JW. A model to accumulate fractionated dose in a deforming organ. *Int J Radiat Oncol Biol Phys* 1999;44:665–675.
28. Nioutsikou E, Webb S, Panakis N, *et al.* Reconsidering the definition of a dose-volume histogram. *Phys Med Biol* 2005;50:L17–L19.

PHYSICS CONTRIBUTION

AUTOMATED CONTOUR MAPPING WITH A REGIONAL DEFORMABLE MODEL

MING CHAO, PH.D., TIANFANG LI, PH.D., EDUARD SCHREIBMANN, PH.D.,
ALBERT KOONG, M.D., AND LEI XING, PH.D.

Department of Radiation Oncology, Stanford University School of Medicine, Stanford, CA

Purpose: To develop a regional narrow-band algorithm to auto-propagate the contour surface of a region of interest (ROI) from one phase to other phases of four-dimensional computed tomography (4D-CT).

Methods and Materials: The ROI contours were manually delineated on a selected phase of 4D-CT. A narrow band encompassing the ROI boundary was created on the image and used as a compact representation of the ROI surface. A BSpline deformable registration was performed to map the band to other phases. A Mattes mutual information was used as the metric function, and the limited memory Broyden-Fletcher-Goldfarb-Shanno algorithm was used to optimize the function. After registration the deformation field was extracted and used to transform the manual contours to other phases. Bidirectional contour mapping was introduced to evaluate the proposed technique. The new algorithm was tested on synthetic images and applied to 4D-CT images of 4 thoracic patients and a head-and-neck Cone-beam CT case.

Results: Application of the algorithm to synthetic images and Cone-beam CT images indicates that an accuracy of 1.0 mm is achievable and that 4D-CT images show a spatial accuracy better than 1.5 mm for ROI mappings between adjacent phases, and 3 mm in opposite-phase mapping. Compared with whole image-based calculations, the computation was an order of magnitude more efficient, in addition to the much-reduced computer memory consumption.

Conclusions: A narrow-band model is an efficient way for contour mapping and should find widespread application in future 4D treatment planning. © 2008 Elsevier Inc.

Deformable model, Image registration, Contour mapping, IGRT.

INTRODUCTION

Segmentation of a region of interest (ROI), such as a tumor target volume or a sensitive structure, is an important but time-consuming task in radiotherapy (1–6). With the emergence of four-dimensional (4D) imaging and adaptive radiotherapy, the need for efficient and robust segmentation tools is even increasing (7–11). Because of dramatically increased numbers of images, it becomes impractical to manually segment the ROIs slice by slice as in current three-dimensional radiotherapy practice. A natural solution to the 4D computed tomography (4D-CT) segmentation problem is to delineate the ROIs on a selected phase and then propagate the contours onto other phases using a mathematical model. Along this line, deformable model-based contour mapping has been implemented by a few groups (12–14). Although feasible, the calculation is global in nature and thus computationally intensive. In addition, the accuracy of the mapped contours may be compromised because the registration may be

influenced unnecessarily by the image content distant from the ROIs, which would otherwise be irrelevant to the contour mapping process. This is especially problematic when non-local deformable models, such as thin plate spline and elastic model, are used. In general, contour mapping is a regional problem, and a global association of the phase-based images is neither necessary nor efficient.

Surface mapping techniques (15–17) represent a competitive alternative to the deformable model-based approach. The idea of surface mapping is to obtain contour transformation by iteratively deforming the ROI contour-extended surface until the optimal match with the reference is found. The calculation involves only the surface region and is thus computationally efficient. Numerous surface mapping techniques have been developed in the past, which include, to name a few, spatial partitioning, principal component analysis, conformal mapping, rigid affine transformation, deformable contours, and warping based on the thin-plate spline. All of

Reprint requests to: Lei Xing, Ph.D., Department of Radiation Oncology, Stanford University School of Medicine, 875 Blake Wilbur Drive, Stanford, CA 94305-5847. Tel: (650) 498-7896; Fax: (650) 498-4015; E-mail: lei@reyes.stanford.edu

Supported in part by grants from the Department of Defense (W81XWH-06-1-0235 and W81XWH-05-1-0041), Komen Breast Cancer Foundation (BCTR0504071), and National Cancer Institute

(5R01 CA98523 and 1 R01 CA98523).

Conflict of interest: none.

Acknowledgment—The authors thank Dr. B. Loo from Stanford University for useful discussions.

Received March 12, 2007, and in revised form Sept 18, 2007.

Accepted for publication Sept 27, 2007.

these techniques are a mapping between topologic components of the input surfaces that allow for transfer of annotations. Although the calculations are inherently efficient, the results depend heavily on the model used, which may not be generally applicable for all clinical situations because the ROI surface is multidimensional and hardly modeled by only a few parameters.

In this work, we present a novel regional algorithm for ROI propagation among different 4D-CT phases. The deformation of an ROI contour-extended surface in our algorithm is not driven by an *ad hoc* surface-based model but instead by the image features in the neighborhood of the surface. The underlying hypothesis here is that information contained in the ROI boundary region is sufficient to guide the contour mapping process. In the proposed algorithm the neighborhood image features of an ROI are captured by a narrow band, which is composed of all points within two surfaces with the signed distances of $\pm d$ from the ROI boundary. The algorithm is a hybrid of the regional surface-based model and the global deformable registration-based approach. The combination takes advantage of the desirable features of each of these two techniques and provides a robust and computationally efficient contour propagation tool for 4D radiotherapy.

METHODS AND MATERIALS

Software platform

The proposed contour mapping algorithm was implemented using the Insight Toolkit (18) and the Visualization Toolkit (19), which are open source cross-platform C++ software toolkits sponsored by the National Library of Medicine.

Overview of the mapping process

Figure 1 depicts the overall contour mapping process. For a given 4D-CT image set, a selected phase, named the template phase, was selected, and the ROIs were manually delineated by a physician. The manually outlined contour was referred to as the template contour. A narrow band encompassing the template contour was created (see next section for details). A deformable mapping was then carried out to propagate the band from the template phase to other phases, referred to as target phases. Upon successful mapping of the band, the deformation field was used to transform the template contour to the target images.

Narrow-band representation of ROI contour

The contour manually segmented on an axial slice of the template image has a polygon shape, and the vertices of the polygon form the basis for constructing the narrow band. As schematically shown in Fig. 2, a band with signed distances $\pm d$ was placed along the template contour. The regional image features contained in the band function serve as a “signature” of the contour and drive the contour mapping process. The distance between the neighboring vertices on the contour is typically 2–10 mm, depending on the shape of the contour. In generating the narrow band, we first created cubes with a side length of $2d$ around all the vertices, as depicted by points A and B in Fig. 2. To obtain a smooth band, between A and B three more cubes, centered at points C, D, and E, were inserted. Point C was chosen to be the middle point between A and B, point D the middle between A and C, and point E the middle between C and

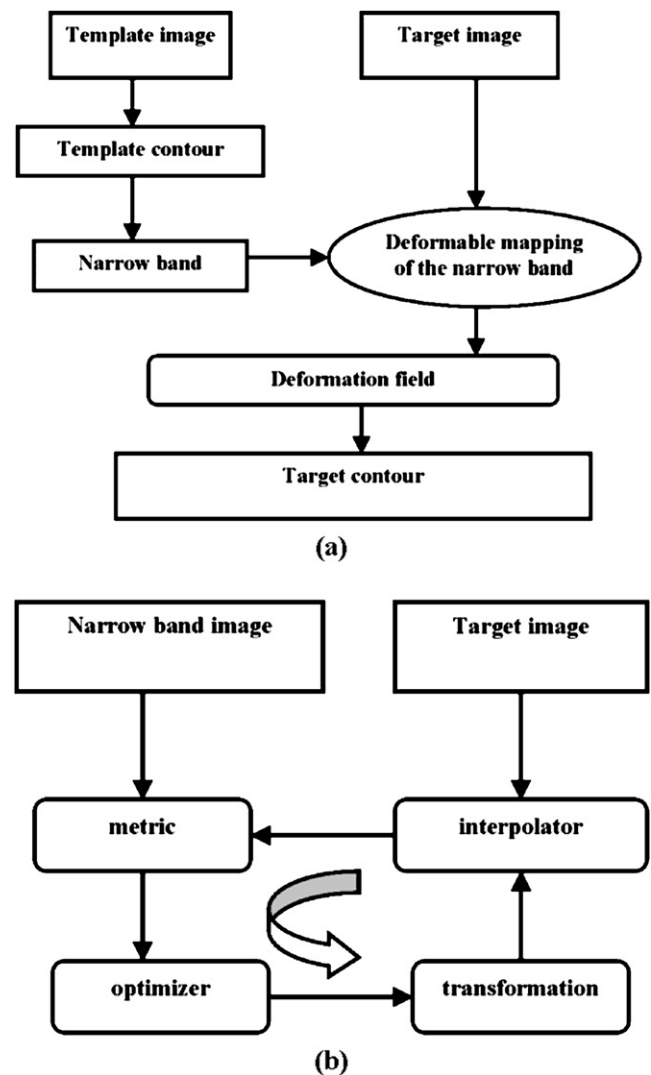


Fig. 1. Flow chart of narrow band-based contour mapping procedure. (a) Overall calculation process. (b) Deformable mapping process of the narrow band.

B. More interpolated vertex points can be introduced similarly when needed. Figure 3 illustrates a narrow band surrounding the lung boundary on the template phase CT image. The light green area stands for the narrow band, and the green curve is the manual contour. The width of the narrow band was set to be $2d = 15$ mm in our calculations. To examine the robustness of the proposed mapping algorithm, a variety of other bandwidths, ranging from 4 mm through 30 mm, were also tested for one of the clinical cases.

Contour propagation

As illustrated in Fig. 1, the process of contour mapping is essentially to warp the narrow band constructed above in such a way that its best match in the target image is found. Mathematically, the mapping process of the narrow band constitutes an optimization problem, in which a group of transformation parameters that transform the points within the band in the template phase to their homologous points in the target image. The warping of the narrow band is quantified by a metric function, which ranks a trial matching based on the “accordance” level of the image content of the band and its correspondence in the target image. The calculation process is detailed below.

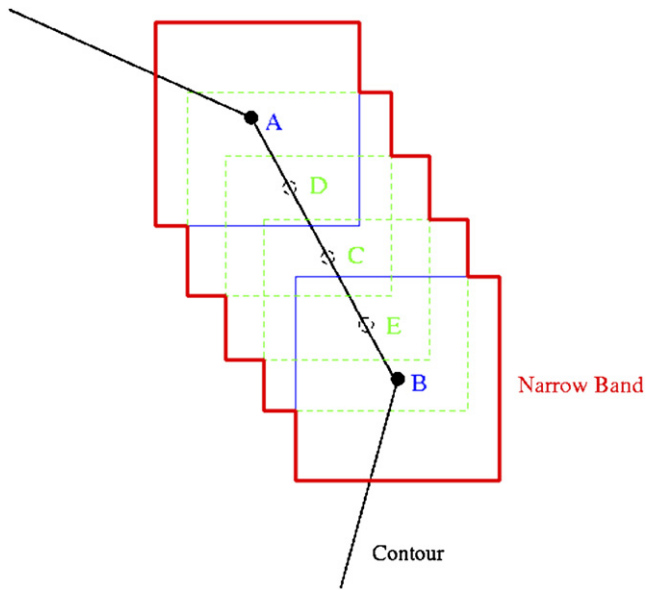


Fig. 2. A schematic drawing of narrow-band construction.

The input to the contour mapping software includes the narrow band and the whole target image, which are described by the image intensity distributions $I_a(\mathbf{x})$ and $I_b(\mathbf{x})$, respectively. It is worth emphasizing that, even though the whole target image was used, only fractional voxels in the target image (the voxels encompassed by the band) are involved in each iteration (a subregion surrounding the ROI on the target image could be created and used in the calculation, but the algorithm converged so fast that after two to three iterations the searching was quickly confined in the neighborhood of the optimal solution). The narrow band acts as a representation of the ROI contour. The task is to find the transformation matrix, $T(\mathbf{x})$, that maps an arbitrary point in the band to the corresponding point on the target image (or *vice versa*) so that the best possible correspondence, as measured by the metric function, is achieved. The calculation proceeds iteratively. A BSpline deformable model is used to model the deformation of the band, but other models should also be applicable. The spacing between the BSpline nodes was chosen to be approximately 0.5 cm (smaller spacing was tested, but no significant difference was found in the final registration results). The displacement of a node i is specified by a vector \mathbf{x}_i , and the displacement vectors (20) of a collection of nodes characterize the tissue deformation. The displacement at a location \mathbf{x} on the image is deduced by a BSpline polynomial fitting.

The Mattes Mutual Information (MMI) (21) was used as the metric function for narrow-band mapping (22–25). The central concept

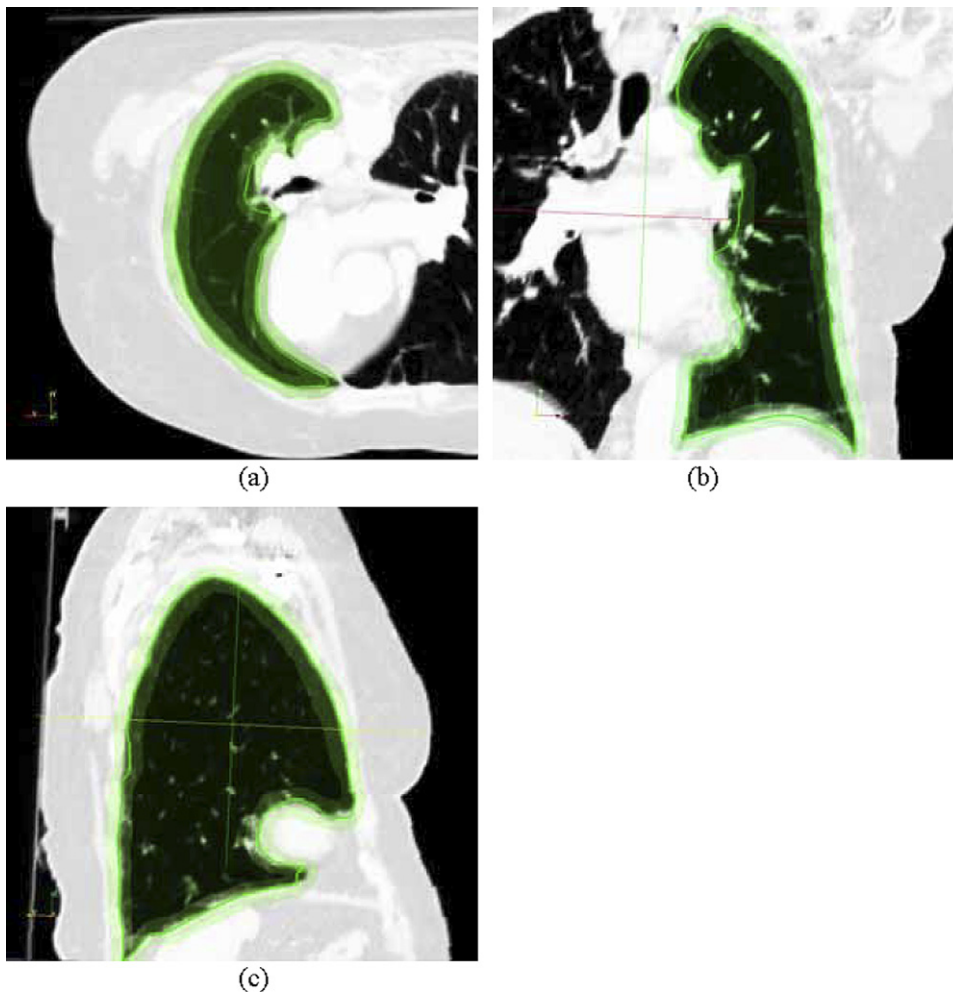


Fig. 3. Computed tomographic images with manual contours and the narrow bands for patient 1. The narrow bands are shown in light green and the contours are green curves. (a) Transverse view; (b) coronal view; (c) sagittal view.

of mutual information (MI) is the calculation of entropy. For an image A, the entropy is defined as

$$H(A) = - \int p_A(a) \log p_A(a) da,$$

where $p_A(a)$ (also called the marginal probability density function [PDF]) is the probability distribution of grey values (image intensities), which is estimated by counting the number of times each grey value occurs in the image and dividing those numbers by the total number of occurrences. Given two images, A and B, their joint entropy is

$$H(A, B) = - \iint p_{AB}(a, b) \log p_{AB}(a, b) dadb,$$

where $p_{AB}(a, b)$ is the joint PDF defined by a ratio between the number of grey values in the joint histogram (feature space) of two images and the total entries (26). The mutual information is generally expressed as

$$MI(A, B) = H(A) + H(B) - H(A, B).$$

Mutual information measures the level of information that a random variable (e.g., $\mathbf{I}_a(\mathbf{x})$) can predict about another random variable (e.g., $\mathbf{I}_b(\mathbf{x})$). Different from the conventional MI, whereby two separate intensity samples are drawn from the image, the Mattes implementation, MMI, uses only one set of intensity to evaluate both the marginal and joint PDFs at discrete positions or bins that uniformly spread within the dynamic range of the images. Entropy values were computed by summing over all the bins. The number of bins used to compute the entropy in MMI metric evaluation was chosen to be 30, and the number of spatial samples used was 20,000. Details of MMI implementation can be found in Mattes *et al.* (21).

The limited memory Broyden-Fletcher-Goldfarb-Shanno algorithm (L-BFGS) (27–29) was used to optimize the MMI metric function with respect to the displacement parameters of the nodes, $\{\mathbf{x}_i\}$, to find the transformation matrix $\mathbf{T}(\mathbf{x})$ that relates the points on image A and image B. Here we just briefly show the algorithm. Starting from a positive definitive approximation of the inverse Hessian \mathbf{H}_0 at \mathbf{x}_0 , L-BFGS derives the optimization variables by iteratively searching through the solution space. At an iteration k , the calculation proceeds as follows: [1] determine the descent direction $\mathbf{p}_k = -\mathbf{H}_k \nabla f(\mathbf{x}_k)$; [2] line search with a step size $\alpha_k = \arg \min_{\alpha \geq 0} f(\mathbf{x}_k + \alpha \mathbf{p}_k)$, where α is the step size defined in the L-BFGS software package; [3] update $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$; and [4] compute \mathbf{H}_{k+1} with the updated \mathbf{H}_k .

At each iteration a backtracking line search is used in L-BFGS to determine the step size of movement to reach the minimum of f along the ray $\mathbf{x}_k + \alpha \mathbf{p}_k$. For convergence α has to be chosen such that a sufficient decrease criterion is satisfied, which depends on the local gradient and function value and is specified in L-BFGS by the Wolfe conditions (27). During the course of optimization, the above iterative calculation based on L-BFGS algorithm continues until the following stopping criterion is fulfilled:

$$\frac{\|\nabla f(\mathbf{x}_k)\|_2}{\max(1, \|\mathbf{x}_k\|_2)} < \varepsilon$$

or a pre-set maximum number of iterations is reached. In this study we set $\varepsilon = 10^6$ and the iteration number to 200, but no more than 100 iterations were exceeded in all our calculations for the algorithm to converge.

Evaluation of algorithm performance

Evaluation of a contour mapping algorithm is a difficult task because of the lack of the ground truth for comparison. A straightforward means of evaluation is the visual inspection of the mapped contours. In addition to this, evaluation based on synthetic images (digital phantoms) is also commonly used. The images and existing contours are distorted with preset deformation fields. Because the gold standard is known, a direct comparison with the mapped contour is made so as to assess the propagation algorithm quantitatively. Beside these two methods, we further performed a bidirectional mapping to evaluate the proposed algorithm. In this test, the reverse of the original contour mapping was performed: the mapped contours on the target phase were treated as the template contours and mapped back to the original template phase. The contours so obtained were then compared with the original manual contours, and the difference between the two sets of contours was quantified. The difference between the resultant and template contours was measured in terms of the displacements of the vertex points on the two contours. The last yet pragmatic evaluation of the algorithm performance on patient's study was based on the physician's manual contours.

Case study

Four thoracic cancer patients, named as patient 1, 2, 3, and 4, were first used to test the proposed algorithm. These patients underwent 4D-CT scans. The 4D-CT images were acquired with a GE Discovery-ST CT scanner (GE Medical System, Milwaukee, WI). The collected data were sorted into 10 phase bins. The ROIs on the template phase were manually segmented by a physician. Specifically, for patients 1 and 2, the inhale phase was chosen for manual segmentation, and for patients 3 and 4, the exhale phase. Different ROIs were used to better evaluate the algorithm. Lungs were selected from patients 1, 2, and 3 and gross tumor volume (GTV) from patient 4. Figure 3 illustrates the manual contour and narrow band representation for the lung from patient 1. Contour is shown in the green curve together with the regional narrow bands (light green area) on the transverse, coronal, and sagittal views (Figs. 3a, 3b, and 3c, respectively).

To further assess the robustness of the proposed algorithm, we also carried out the contour propagation calculation from planning CT to Cone-beam CT (CBCT) for a head-and-neck case. The CBCT images were acquired using the Varian Trilogy system (Varian Medical Systems, Palo Alto, CA).

RESULTS

Convergence analysis

To better illustrate the iterative process of the contour propagation, in Fig. 4 the MMI metric as a function of iteration step is plotted for the narrow band mapping from the first phase (inhale phase) to the other nine phases for the first thoracic patient. In all nine calculations it is seen that the metric value decreases monotonically as the iteration proceeds. However, the number of iterations needed for the algorithm to find the optimal solution varies. It is interesting to observe that, for an “easier” mapping whereby the deformation between the two phases is small, the number of iterations required is less, whereas for “tougher” ones with larger differences in ROI shapes, the required number of iterations increases drastically. Indeed, from Fig. 4 it is seen that the minimum number of iterations required for the metric to saturate occurs when mapping the phase 1 to the adjacent

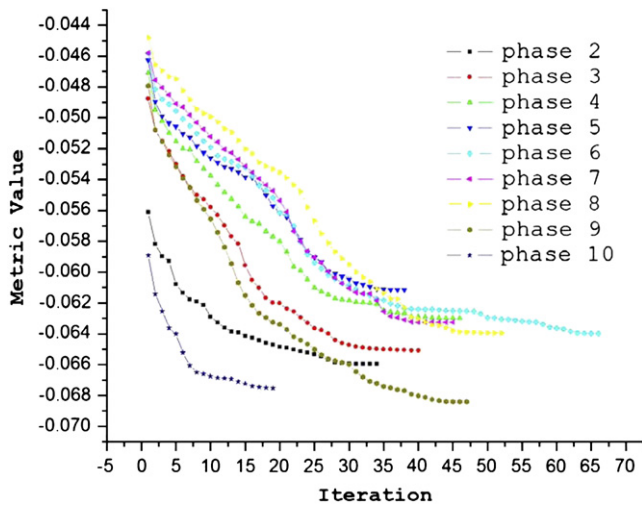


Fig. 4. Narrow-band metric values as a function of iteration step when mapping the narrow band from phase 1 to the other nine phases of the four-dimensional computed tomography.

phases, 2 and 10. For other mapping, the required iteration increases and reaches its largest value for the “toughest” mapping between inhale and exhale (phase 5) phases.

In the above analysis, the bandwidth was set to be 15 mm. The performance of the proposed algorithm was also evaluated by varying the width in the range of 4 mm and 30 mm. Specifically, we tried the widths of 4 mm, 8 mm, 10 mm, 15 mm, 20 mm, and 30 mm. Our results revealed that, when the band was too narrow (e.g., 4 mm), the mapping may fail locally at a place not containing sufficient neighborhood image features. The situation is improved dramatically as the bandwidth increases. For all the clinical cases studied here, no single failure was observed for a width of 15 mm. When the width is too large, the whole ROI will be included in the band. In this situation, the mapping becomes equivalent to registering the whole image and the advantage of the narrow band will be overshadowed by the dramatically increased memory and computing costs. Our experience indicates that a width of 10–15 mm provides a fine balance between the computational accuracy and the associated cost.

We found that the overall computing time was increased by roughly an order of magnitude when going from the narrow band approach to the conventional deformable model-based contour mapping, say, approximately 3 min for narrow band-based mapping vs. approximately 25 min for whole image-based mapping. The dramatically increased computer memory requirement in the latter case also posts a serious problem when developing a clinically practical contour propagation method for 4D radiotherapy.

Algorithm performance evaluation

In addition to visual inspect, the proposed algorithm was assessed by a series of synthetic images or digital phantoms. Typically, a thoracic CT image together with the contour was distorted with the intentionally introduced deformation, and then the contour was propagated onto the distorted image.

A quantitative comparison was carried out. The mean and maximum separation between the gold standard and the mapped contours were found to be 1.0 mm and 1.5 mm, respectively. Figure 5 shows one example of digital phantom experiments.

The performance of the proposed algorithm was further evaluated by the bidirectional mapping calculation outlined in Methods and Materials. A template contour at phase 1 was first mapped to phases 3 and 6. The mapped contours were then treated as the “starting contours” and mapped back to phase 1. The two back-mapped contours were compared with the original template contour. The displacement of each back-mapped vertex point relative to its original locations was computed, and a mean value of 0.8 mm was found for the bidirectional mapping between phases 1 and 3 and 1.8 mm between phases 1 and 6. The larger displacement in the latter situation was due to the fact that, computationally, it is more difficult to map between two opposite phases, such as inhale and exhale phases, owing to larger organ deformations. Overall, the observed displacement is comparable to the pixel size, indicating that the mapping is accurate and robust.

Thoracic patient study results

Figure 6 shows the contour mapping results for the first clinical case. The results are presented in axial, coronal, and sagittal planes for phases 2 (Fig. 6a–c), 6 (Fig. 6d–f), 8 (Fig. 6g–i), and 10 (Fig. 6j–l). For phases 2 and 10, which are immediately adjacent to the inhale phase, the deformation is relatively small and the mapped contours conform to the ROI boundary very well. This represents the “easy” mapping situation and is consistent with the analysis presented above. The average error was less than 1.5 mm. For a “remote” phase, such as phase 6 shown in Fig. 6d–f, more

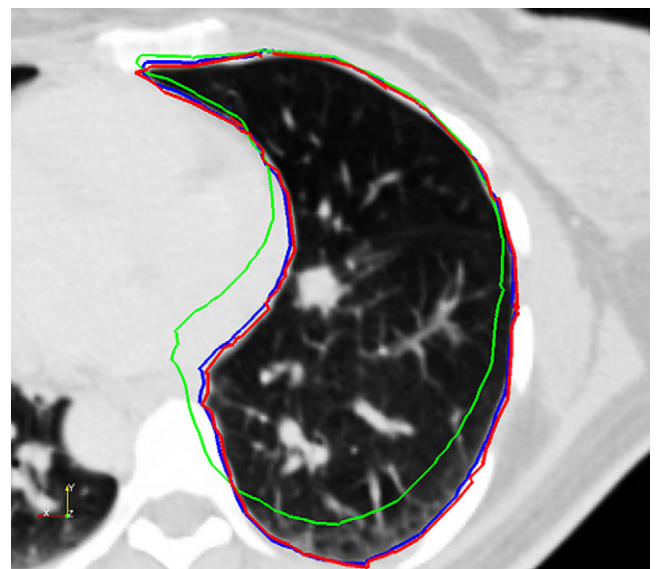


Fig. 5. Synthetic image and overlaid contours. The original contour is depicted in green, gold standard contour in blue, and the mapped contour in red.

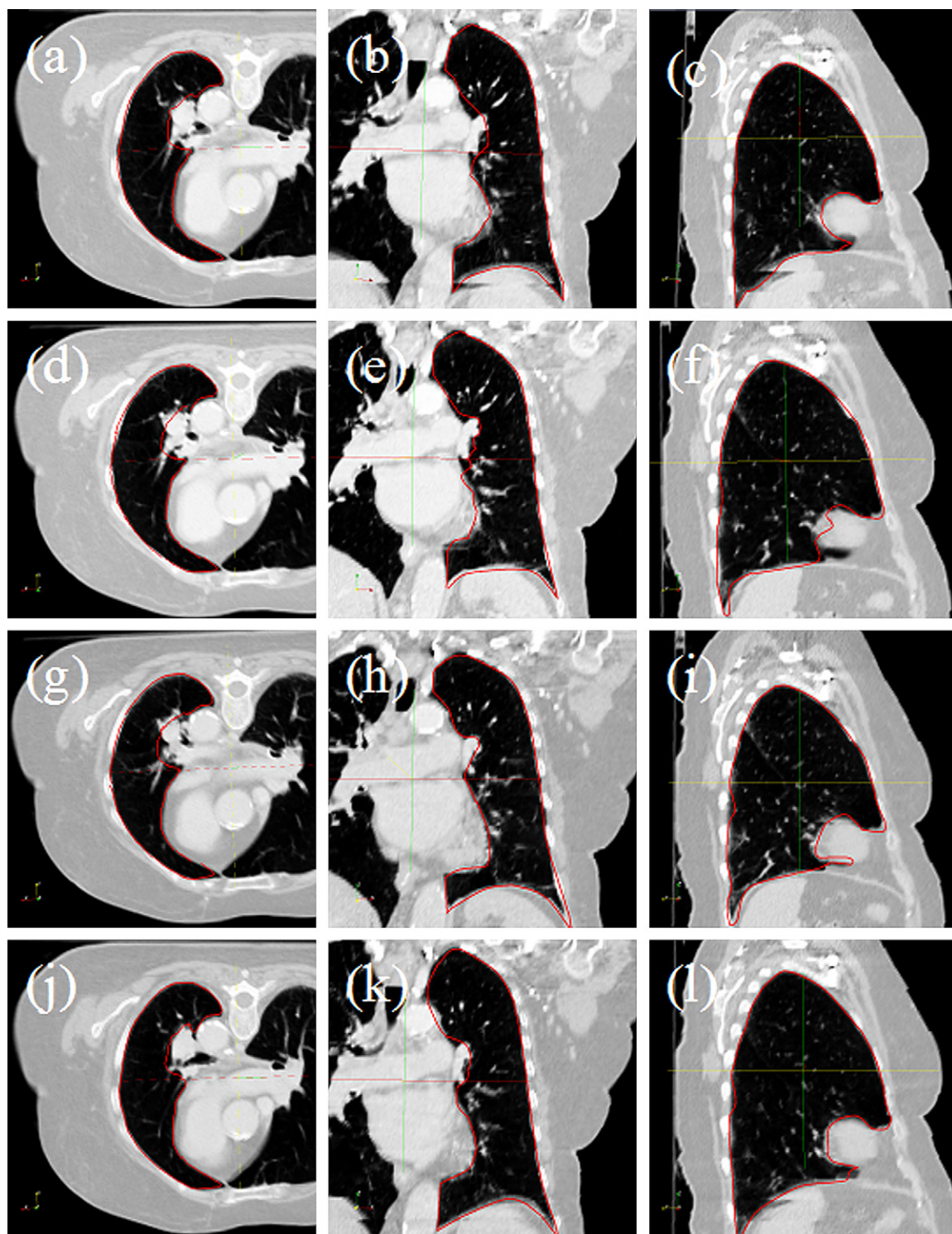


Fig. 6. Computed tomographic images and mapped contours for thoracic patient 1. Displayed are selected phases. From the top row to bottom, phases 2, 6, 8, and 10 are presented, respectively. For each phase, transverse, coronal, and sagittal views are shown from left to right.

iterations were entailed to find the optimal solution, and the resultant contours tend to be worse as compared with those phases adjacent to phase 1. According to the bidirectional mapping, the average mapping error for phase 6 was estimated to be less than 3 mm. The mapped GTV contours (in red) together with manual contours (in blue) by a physician

for phases 1, 4, 8, and 10 in the study of patient 4 are shown in Fig. 7 (parts a, b, d, and e, respectively). The template phase (phase 6) with the template manual contour (in green) is shown in Fig. 7c. In addition, the template manual contour from this phase was overlaid on all the displayed phases. For phases 4 and 8 the deformation was relatively

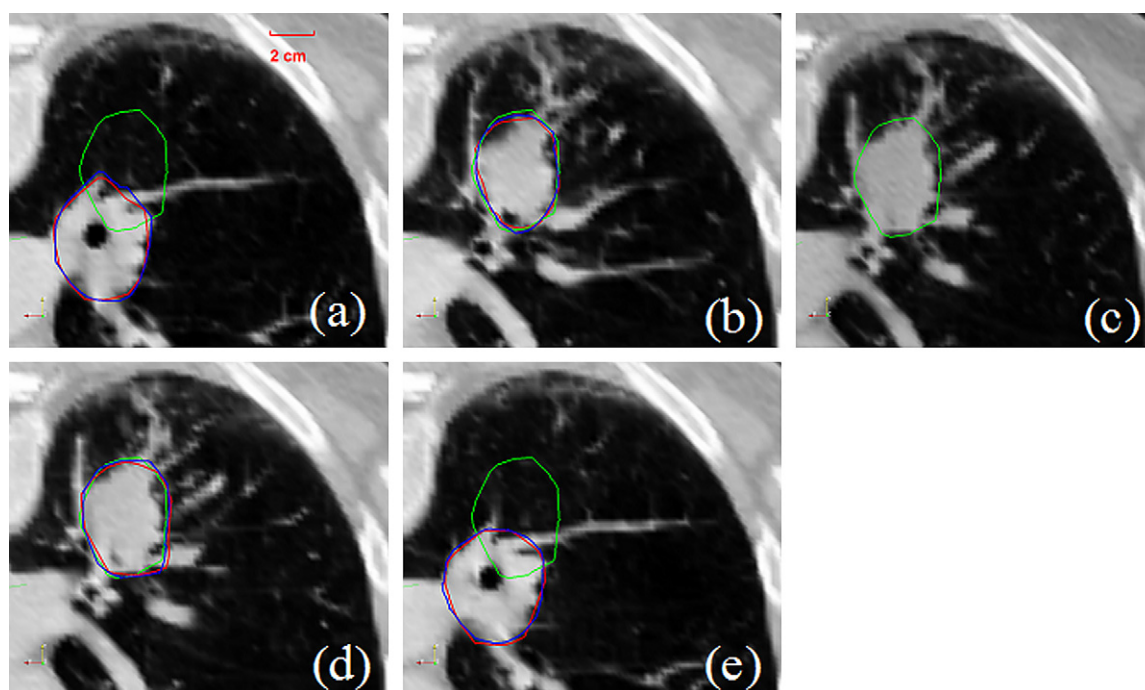


Fig. 7. Axial view of computed tomographic images with gross tumor volume contours for the fourth thoracic patient. (a), (b), (c), (d), and (e) correspond to phases 1, 4, 6 (template phase), 8, and 10, respectively. The green curves are the manually outlined template contour from phase 6, and the red curves represent the contours after warping. The manual contours (in blue) by a physician on individual phases were also displayed.

small (the manual contour was delineated on phase 6), and fewer iterations were needed to find the optimal bands on the target images. For phases 1 and 10, whereby deformation was significant in the ROIs although more computing load was necessary, a good result was still achieved with our narrow-band technique. Comparisons between the mapped contours and the manually segmented contours by physicians for these patients were also performed, and results revealed a similar level of accuracy (maximum and mean values of the discrepancy between the two sets of contours are 2.8 mm and -0.9 mm, respectively).

As a useful application of the proposed technique, in Fig. 8 we present the mean and maximum lung displacements of contour vortices for each breathing phase relative to their locations on the template phase. As seen in Fig. 8, the overall behavior of the mean and maximum displacements is consistent with our intuitive expectation. For cases 1 and 2, the inhale phase (phase 1) was manually segmented, thus the displacement for that phase is zero. For other phases, both mean and maximum displacement values vary with the breathing phase and reach their maxima at the opposite phase. For case 3 the exhale phase was manually segmented, and the behavior was thus opposite to cases 1 and 2. In general, an average displacement of approximately 3 mm was found for inhale and exhale phases. A slight digression is noticed in phase 7 of patient 1, which may be caused by 4D-CT binning artifacts. This type of data is particularly useful in determining the patient-specific tumor margin to account for breathing motion of the tumor target.

Contour propagation in a head-and-neck case

The results of contour mapping for the head-and-neck case are summarized in Fig. 9. Figure 9a shows the planning CT along with manually delineated contours, and Fig. 9b displays the mapped contours of the body, mandible, and GTV on CBCT. For body and mandible a simple rigid mapping is enough to achieve high accuracy. For the GTV, however, the proposed deformable registration model was necessary to adequately propagate the contour. A visual inspection of the propagated contours suggests that the mapping is clinically acceptable.

DISCUSSION

Four-dimensional CT image segmentation represents a necessary step in constructing a 4D patient model and computing the accumulated dose in 4D radiotherapy. A natural way to tackle the problem is to auto-map the manually delineated contours on one of the phases to the remaining phases. In this work, a regional computing algorithm was introduced to deal with the issue. The approach relies on the assumption that a narrow band surrounding the manually segmented contour can capture sufficient information to drive the finding of its counterparts in other phases of the 4D-CT. Obviously, this assumption is valid when the band is sufficiently wide so that a large number of voxels are involved in the registration calculation. As demonstrated by the presented data, the registration and the mapping are reliable when the bandwidth is larger than 4 mm. Computationally, the proposed approach

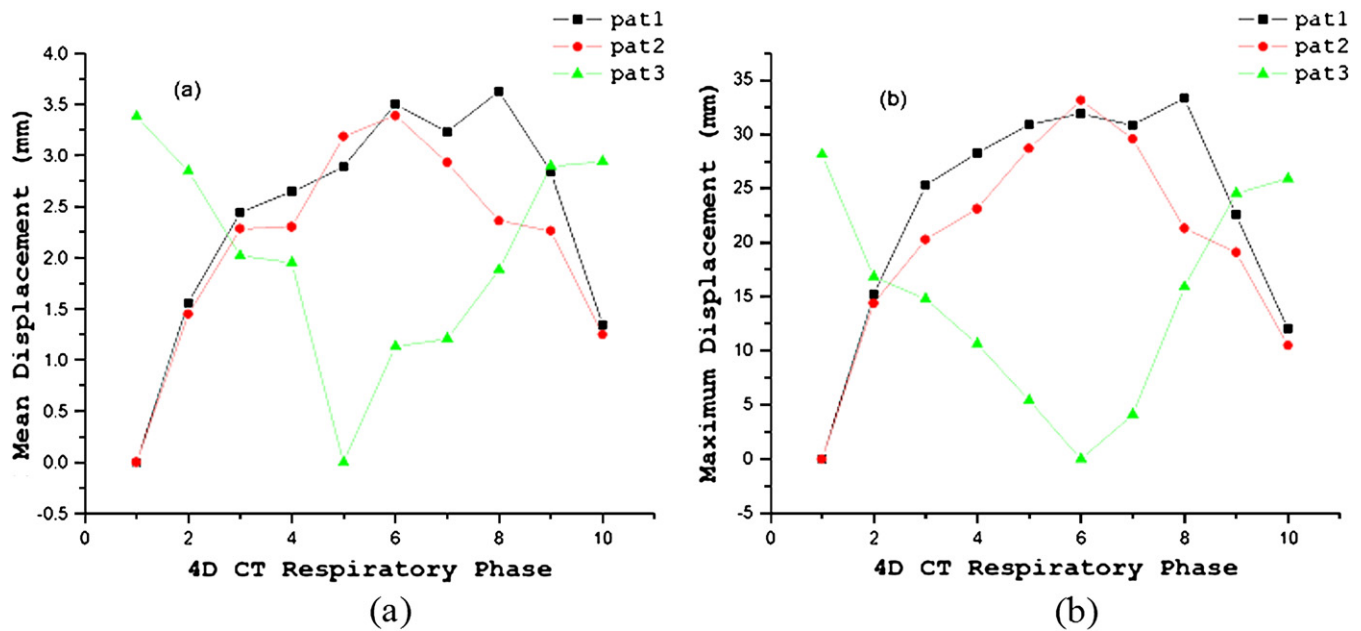


Fig. 8. Displacement of region of interest boundary points as a function of respiration phase for three thoracic patients. (a) Mean displacement vs. phase. (b) Maximum displacement vs. phase. 4D CT = four-dimensional computed tomography.

resides between a deformable model-based mapping and a surface model-based ROI contour mapping.

The success of the image content-based approaches, such as the proposed narrow-band approach or conventional deformable image registration, arises from the fact that they fully utilize the inherent image features of the two input images. The narrow band-based technique is particularly attractive because it takes advantages of the useful features of both image content-based technique and the regional surface-based model. In a sense, it is a hybrid approach of the two distinct types of algorithms. The narrow-band approach utilizes the imaging features surrounding the ROI to guide the search of the optimal mapped contours while considering the shape integrity of the ROI surface. It eliminates the need for a global

registration of the input images and thus greatly increases the computational efficiency.

Application of the proposed contour mapping technique to five clinical cases indicates that the technique is accurate and computationally efficient. A common problem in image segmentation and contour mapping studies is the lack of quantitative validation. In the studies of Lu *et al.* (13) and Schrieblmann *et al.* (14), for example, the accuracy of a deformable model-based contour mapping technique was evaluated purely on the basis of visual inspection. Although it is a convenient way for rapid assessment of a segmentation calculation, especially in a case in which the “ground truth” contours do not exist, the method falls short in quantization. The same approach was used in many other previous

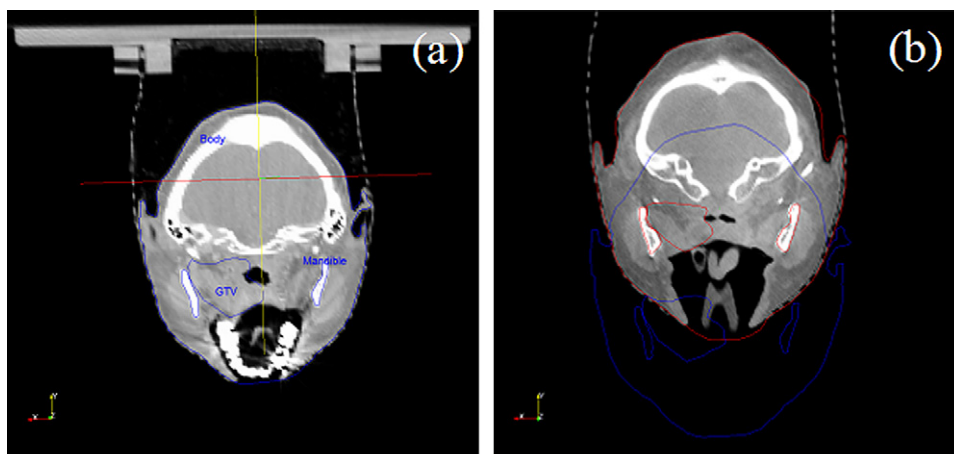


Fig. 9. Contour propagation in a head-and-neck case. (a) Planning computed tomography with manually outlined template contours (in blue) for body, mandible, and gross tumor volume. (b) Cone-beam computed tomography along with contours after warping (in red) for the corresponding structures.

investigations (1, 5, 14, 30). In this study, a bidirectional contour mapping was proposed to examine the reliability and robustness of a contour mapping technique. This method provides a useful test in assessing the success of a contour propagation algorithm. We would like to point out that the bidirectional mapping technique introduced in this work is a necessary (but not sufficient) test. In a rare but possible situation, the bidirectional mapping may not be able to find that an error occurred in the narrow-band mapping process. A visual inspection of the mapped result may help in this situation. On the basis of the bidirectional mapping experiments and visual inspection for the patient studies, we conclude that the proposed approach can perform very well even in the presence of significant deformations.

In our calculation, we observed that the regular grid of BSpline control points could be mapped to a region outside the narrow band. Although it seems that this does not directly affect the accuracy of the method, it may prolong the calculation by computing the displacements in regions where metric information is irrelevant. Setups have been proposed to adapt the splines control mesh to regions where deformation is found to be significant (31), and the extension of the method would allow us to use the BSpline control points defined only in the regions within the narrow band. Implementation of this type of technique should further reduce the computation time required to find the optimal solution.

Although there are numerous deformable algorithms, including, for example, the elastic model (32–34), viscous fluid model (35), optical flow model (5,30,36), finite element model (33, 37), and radial basis function models such as the basis spline model (28, 38, 39) and thin plate spline model (40–43), a truly robust tool suitable for routine clinical appli-

cations is yet to be developed. Each of these approaches has its pros and cons. The deformable calculation can be greatly facilitated if some *a priori* system information can be incorporated. Along this line, the homologous correspondence of the bony structure in two input images has been incorporated in thin plate spline method, and remarkable improvement has resulted (44). The narrow band-generated ROI contour correspondence could also be used as prior knowledge to improve a deformable registration. This work is still in progress and will be reported in the future.

CONCLUSIONS

In this work we have developed a regional deformable registration-based method to auto-propagate contours for 4D radiotherapy. The central idea is that a narrow band encompassing an ROI surface carries the neighborhood information of the ROI surface and can be used to establish a reliable association between the ROIs in two phase-specific image sets. Different from other type of regional algorithms, such as surface mapping, the method uses the image features captured in a band to guide the search for the optimal contour mapping. Compared with conventional deformable image registration-based approaches, a great reduction in computational burden and a large capture radius in optimization space result. Our study demonstrated that the information contained in the boundary region can be used to guide the contour mapping in all the testing cases presented in this article. The proposed regional model decreases the workload involved in 4D-CT ROI segmentation and provides a valuable tool for the efficient use of available spatial-temporal information for 4D simulation and treatment planning.

REFERENCES

1. Pekar V, McNutt TR, Kaus MR. Automated model-based organ delineation for radiotherapy planning in prostatic region. *Int J Radiat Oncol Biol Phys* 2005;60:973–980.
2. Kass MR, WitKen A, Terzopoulos D. Snakes: Active contour models. *Int J Comput Vis* 1988;4:321–331.
3. Coote T, Hill A, Taylor C, *et al.* The use of active shape models for locating structures in medical images. *J Image Vis Comput* 1994;12:355–366.
4. Xu C, Prince JL. Snakes, shapes, and gradient vector flow. *IEEE Trans Image Process* 1998;7:359–369.
5. Liu F, Zhao B, Kijewski PK, *et al.* Liver segmentation for CT images using GVF snake. *Med Phys* 2005;32:3699–3706.
6. Weese J, Kaus MR, Lorenz C, *et al.* Shape constrained deformable models for 3D medical image segmentation. In: Insana MF, Leahy RM. Information processing in medical imaging: 17th International Conference, IPMI 2001, Davis, CA, USA, June 18–22, 2001. *Lecture Notes in Computer Science* 2001; 2082:380–387.
7. Li T, Schreiber E, Thorndyke B, *et al.* Radiation dose reduction in four-dimensional computed tomography. *Med Phys* 2005;32:3650–3660.
8. Vedam SS, Keall PJ, Kini VR, *et al.* Acquiring a four-dimensional computed tomography dataset using an external respiratory signal. *Phys Med Biol* 2003;48:45–62.
9. Dietrich L, Jetter S, Tucking T, *et al.* Linac-integrated 4D cone beam CT: First experimental results. *Phys Med Biol* 2006;51: 2939–2952.
10. Li T, Xing L, Munro P, *et al.* Four-dimensional cone-beam computed tomography using an on-board imager. *Med Phys* 2006;33:3825–3833.
11. Rietzel E, Chen GT, Choi NC, *et al.* Four-dimensional image-based treatment planning: Target volume segmentation and dose calculation in the presence of respiratory motion. *Int J Radiat Oncol Biol Phys* 2005;61:1535–1550.
12. Gao S, Zhang L, Wang H, *et al.* A deformable image registration method to handle distended rectums in prostate cancer radiotherapy. *Med Phys* 2006;33:3304–3312.
13. Lu W, Olivera GH, Chen Q, *et al.* Automatic re-contouring in 4D radiotherapy. *Phys Med Biol* 2006;51:1077–1099.
14. Schreiber E, Chen GT, Xing L. Image interpolation in 4D CT using a BSpline deformable registration model. *Int J Radiat Oncol Biol Phys* 2006;64:1537–1550.
15. Chakraborty A, Staib LH, Duncan JS. An integrated approach for surface finding in medical images. In: IEEE workshop mathematical methods in biomedical image analysis. Los Alamitos, CA: IEEE Computer Society Press; 1996. p. 253–262.
16. McNerney T, Terzopoulos D. Deformable models in medical image analysis. *Med Image Anal* 1996;1:91–108.

17. Montagnat J, Delingette H, Ayache N. A review of deformable surfaces: Topology, geometry and deformation. *Image Vis Comput* 2001;19:1023–1040.
18. Ibanez L, Schroeder W, Ng L. ITK software guide. Clifton Park, NY: Kitware; 2003.
19. Schroeder W, Martin K, Lorensen B. The visualization toolkit: An object-oriented approach to 3D graphics. 4th edition. Kitware: Clifton Park, NY; 2006.
20. Chao M, Schreibmann E, Li T, *et al.* Knowledge-based auto-contouring in 4D radiation therapy. *Med Phys* 2006;33:2171.
21. Mattes D, Haynor DR, Vesselle H, *et al.* Non-rigid multi-modality image registration. In: Sonka M, Hanson KM, editors. *Medical imaging 2001: Image processing. Proceedings of SPIE* 2001;4322:1609–1620.
22. Woods RP, Cherry SR, Mazziotta JC. Rapid automated algorithm for aligning and reslicing PET image. *J Comput Assist Tomogr* 1992;16:620–633.
23. Woods RP, Mazziotta JC, Cherry SR. MRI-PET registration with automated algorithm. *J Comput Assist Tomogr* 1993;17:536–546.
24. Collingnon A, Maes F, Delaere D, *et al.* Automated multi-modality image registration based on information theory. In: Bizais Y, Barillot C, Paola RD, editors. *Information processing in medical imaging*. Dordrecht, The Netherlands: Kluwer; 1995. p. 263–274.
25. Wells WM III, Viola P, Kikinis R. Multi-modal volume registration by maximization of mutual information. *Med Image Anal* 1996;1:35–51.
26. Pluim JP, Maintz JB, Viergever MA. Mutual-information-based registration of medical images: A survey. *IEEE Trans Med Imaging* 2003;22:986–1004.
27. Liu DC, Nocedal J. On the limited memory BFGS method for large scale optimization. *Math Program* 1989;45:503–528.
28. Schreibmann E, Yang Y, Boyer A, *et al.* Image interpolation in 4D CT using a BSpline deformable registration model. *Med Phys* 2005;32:1924.
29. Schreibmann E, Xing L. Narrow band deformable registration of prostate magnetic resonance imaging, magnetic resonance spectroscopic imaging, and computed tomography studies. *Int J Radiat Oncol Biol Phys* 2005;62:595–605.
30. Guerrero TM, Zhang G, Huang TC, *et al.* Intrathoracic tumour motion estimation from CT imaging using the 3D optical flow method. *Phys Med Biol* 2004;49:4147–4161.
31. Camara O, Colliot O, Delso G, *et al.* 3D nonlinear PET-CT image registration algorithm with constrained free-form deformations. In: Hamza MH, editor. *Proceedings of the 3rd IASTED International Conference on Visualization, Imaging, and Image Processing*. Calgary: ACTA Press; 2003. p. 516–521.
32. Bajcsy R, Kovacic S. Multiresolution elastic matching. *Comput Vis Graphics Image Processing* 1989;46:1–21.
33. Gee JC, Haynor DR, Reivich M, *et al.* Finite element approach to warping of brain images. *Proc SPIE Med Imaging* 1994;2167:18–27.
34. Gee JC, Reivich M, Bajcsy R. Elastically deforming 3D atlas to match anatomical brain images. *J Comput Assist Tomogr* 1993;17:225–236.
35. Christensen GE, Rabitt RD, Miller MI. Deformable templates using large deformable kinematics. *IEEE Trans Med Imaging* 1996;5:1435–1447.
36. Thirion JP. Image matching as the diffusion process: An analogy with Maxwell's demons. *Med Image Anal* 1998;2:243–260.
37. Brock KM, Balter JM, Dawson LA, *et al.* Automated generation of a four-dimensional model of the liver using warping and mutual information. *Med Phys* 2003;30:1128–1133.
38. Schreibmann E, Xing L. Image registration with auto-mapped control volumes. *Med Phys* 2006;33:1165–1179.
39. Coselmon MM, Balter JM, McShan DL, *et al.* Mutual information based CT registration of the lung at exhale and inhale breathing states using thin-plate splines. *Med Phys* 2004;31:2942.
40. Lian J, Xing L, Hunjan S, *et al.* Mapping of the prostate in endorectal coil-based MRI/MRSI and CT: A deformable registration and validation study. *Med Phys* 2004;31:3087–3094.
41. Brock KK, Hollister SJ, Dawson LA, *et al.* Technical note: creating a four-dimensional model of the liver using finite element analysis. *Med Phys* 2002;29:1403–1405.
42. Fei B, Kemper C, Wilson DL. A comparative study of warping and rigid body registration for the prostate and pelvic MR volumes. *Comput Med Imaging Graph* 2003;4:267–281.
43. Bookstein FL. Principal warping: Thin plate splines and the decomposition of deformations. *IEEE Trans Pattern Anal Machine Intelligence* 1989;11:567–585.
44. Xie Y, Xing L. Incorporating a priori knowledge into deformable registration model [Abstract]. *Med Phys* 2007;34:2333–2334.

Development of a QA phantom and automated analysis tool for geometric quality assurance of on-board MV and kV x-ray imaging systems

Weihua Mao, Louis Lee, and Lei Xing^{a)}

Department of Radiation Oncology, Stanford University School of Medicine, Stanford, California 94305

(Received 27 September 2007; revised 15 January 2008; accepted for publication 17 January 2008; published 19 March 2008)

The medical linear accelerator (linac) integrated with a kilovoltage (kV) flat-panel imager has been emerging as an important piece of equipment for image-guided radiation therapy. Due to the sagging of the linac head and the flexing of the robotic arms that mount the x-ray tube and flat-panel detector, geometric nonidealities generally exist in the imaging geometry no matter whether it is for the two-dimensional projection image or three-dimensional cone-beam computed tomography. Normally, the geometric parameters are established during the commissioning and incorporated in correction software in respective image formation or reconstruction. A prudent use of an on-board imaging system necessitates a routine surveillance of the geometric accuracy of the system like the position of the x-ray source, imager position and orientation, isocenter, rotation trajectory, and source-to-imager distance. Here we describe a purposely built phantom and a data analysis software for monitoring these important parameters of the system in an efficient and automated way. The developed tool works equally well for the megavoltage (MV) electronic portal imaging device and hence allows us to measure the coincidence of the isocenters of the MV and kV beams of the linac. This QA tool can detect an angular uncertainty of 0.1° of the x-ray source. For spatial uncertainties, such as the source position, the imager position, or the kV/MV isocenter misalignment, the demonstrated accuracy of this tool was better than 1.6 mm. The developed tool provides us with a simple, robust, and objective way to probe and monitor the geometric status of an imaging system in a fully automatic process and facilitate routine QA workflow in a clinic. © 2008 American Association of Physicists in Medicine. [DOI: [10.1118/1.2885719](https://doi.org/10.1118/1.2885719)]

I. INTRODUCTION

The integration of on-board kilovoltage (kV) flat-panel imager to a medical linear accelerator has recently been realized by linac vendors¹⁻⁴ for image-guided radiation therapy (IGRT). These on-board imagers are mounted on robotic arms with an axis orthogonal to the megavoltage beam.⁵ An on-board imager offers three modes of acquisitions namely: (1) Two-dimensional (2D) planar projection image (radiographic acquisition); (2) 2D planar fluoroscopic image (fluoroscopic acquisition); and (3) three-dimensional (3D) cone-beam computed tomography (CBCT). The radiographic acquisition is used for 2D-2D matching with the digital reconstructed radiograph (DRR) for setup verification based on the bony landmarks or implanted fiducials.^{1,6} The fluoroscopic acquisition is employed for verifying the gating threshold or target position prior to a respiratory-gated treatment.⁷ The CBCT is used for 3D-3D matching with the planning CT for setup verification.⁸⁻¹¹ The geometric information of a patient derived from any of these modes relies on the configuration of the imager geometry, such as the position of the x-ray source, imager position and orientation, isocenter, and focus-to-imager distance. The functionality of the system depends heavily on the mechanical integrity and stability of the imaging device (the x-ray tube, imager and robotic arms) and the linac (on which the imaging system is mounted) at various gantry angles. This is particularly crucial for the flat-panel based CBCT that is mounted on an already laden gantry of a linac. The gantry sagging, together

with the flexing of the imaging system, results in geometric nonidealities during the gantry rotation in CBCT acquisition.^{10,12} The accurate reconstruction of transaxial slices of a 3D object from a set of 2D projections requires the x-ray source position and the detector orientation to be known precisely in 3D space in the rotation trajectory during the CBCT reconstruction.^{10,13-15} These geometric parameters are generally established through a geometrical calibration process done during the commissioning of the new machine and incorporated in image formation or reconstruction for correction.^{10,16} A prudent use of an on-board imaging system necessitates a routine surveillance of geometric accuracy of the system. This echoes to what Yoo *et al.*¹⁷ have pointed out that the most crucial part of a comprehensive QA program on an on-board imager would be those tests monitoring the geometric accuracy and stability of the imaging system. The recommendation by Yoo *et al.*¹⁷ on these tests is based on imaging a small cube phantom embedded with a central fiducial at orthogonal gantry angles and manually measuring the discrepancy between the recorded position of the fiducial and the digital graticule. This method is straightforward yet subjective due to the manual measurement. Furthermore, it is unable to give a full picture of the geometric status of the imaging system.

In research involving micro-CBCT, several geometric phantoms and associated analysis software have been developed for the geometric calibration and QA.^{14,18} A modification and an extension of its methodology might be warranted for the geometric QA of the on-board imaging system. Stud-

ies have also reported on the use of projection images of simple phantoms embedded with multiple fiducials to estimate the geometric parameters of the on-board imaging system,^{12,19} however the geometric parameters that can be estimated implicitly from the projection images have not been fully extracted due to the limitations in the design of the phantoms or associated data analysis software.

The purpose of the present work is to design a geometric QA phantom and develop an automated data analysis software to assess the geometric accuracy of an on-board imager, including the x-ray source position, the detector position and orientation, the isocenter, the rotation trajectory, and the source-to-imager distance (SID), from the projection images of the phantom. The developed phantom and analysis software apply equally well to the geometric QA of the MV electronic portal imaging device (EPID), and allow us to measure the coincidence of the isocenters of the MV and kV beams, which is an important parameter in ensuring that the treatment beam is targeting to where the imaging is guiding. The geometric QA procedure using this phantom, together with the automation of the data analysis, greatly facilitates the QA workflow and eliminates the subjectivity incurred in the manual measurement and result interpretation.

II. METHODS AND MATERIALS

II.A. Phantom design

An in-house software developed with the MatLab (MathWorks, Inc., Natick, MA) was used to simulate the projection image of an object with designated imaging geometry defined by the position of the source, position and orientation of the imager, and source-to-imager distance at any gantry angle. Virtual cubic phantoms (instead of physical phantoms) of various dimensions from 10 to 25 cm with 9, 13, 17 fiducials arranged in a helical trajectory at the surfaces were designed. These virtual phantoms served as inputs into the simulation program to simulate the projection images obtained at different gantry angles. The helical arrangement of the fiducials was chosen because it gave a good distribution of the fiducials in 3D space and offered the least superposition of the fiducials in the projection images obtained from different gantry angles for easy detection and identification. Nonuniform helical increments (itches) along the helical axis were used to further spare the fiducial projections.

Simulated projection images were created for each virtual phantom at every 30° of gantry angles. The phantom design, including the size and fiducial locations, was determined by maximizing the detection efficiency and sensitivity of fiducials to a change in imaging geometry. Based on the findings from the simulation, a cubic phantom measuring 18 × 18 × 18 cm³ with 13 steel ball bearings (BBs) of diameter 4.76 mm was fabricated. The BBs were embedded in the surface of the phantom in a skewed helical trajectory as shown in Fig. 1. The outside surface of every BB was tangential to the surface of the phantom ensuring that all BBs were protected from possible displacement or wearing during setup or handling. Orthogonal lines were scribed on the sur-

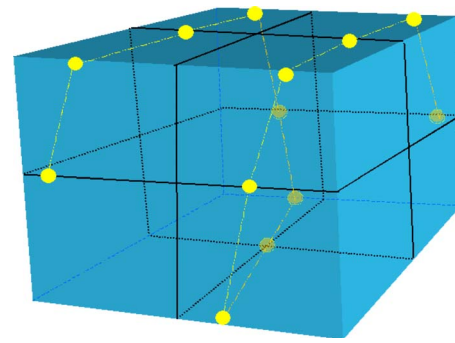


FIG. 1. The *gQA* phantom with 13 steel ball bearings (BBs) mounted on the surfaces in a skewed helical trajectory.

faces of the phantom to indicate the center of the cube at the intersection of these lines. This phantom was termed as *gQA* phantom hereafter for easy reference.

II.B. Reference coordinate system

To describe the geometry of the imaging system, a Cartesian phantom coordinate system was introduced. This coordinate system is attached to the phantom and aligned in space to an ideal linac with a rotating gantry. The *z* axis is along the rotation axis of the gantry, the *x* axis is horizontal, and *y* axis is vertical [Fig. 2(a)]. The origin is referred to the intersection of the laser alignments.

For a gantry angle ϕ (which also indicates the x-ray source angle), the position of the x-ray source (R, ϕ, z) is conveniently defined in a phantom fixed cylindrical coordinate which has the same origin and *z* axis of the phantom coordinate system, where *R* is the radius of the source rotation circle or the source-to-axis distance (SAD). A fiducial at (r, θ, z_B) is projected on the imager (*u, v*) [Fig. 2(b)] with

$$u = \frac{-Fr \cos(\theta - \phi)}{R + r \sin(\theta - \phi)}, \quad (1)$$

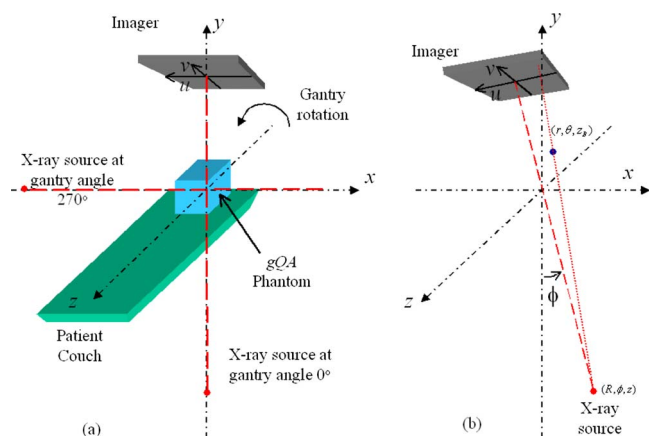


FIG. 2. Schematic representation of the phantom coordinate system (a), and the position of the x-ray source and a fiducial in the cylindrical coordinate system (b).

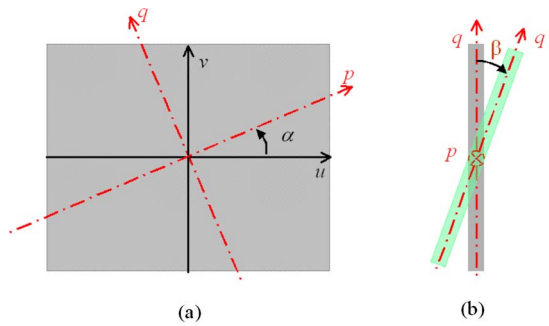


FIG. 3. Schematic diagram showing the tilting of the imager. The imager (uv plane) tilts along an imaginary axis p (a) and the tilting angle is defined as β (b).

$$v = \frac{F(z_B - z)}{R + r \sin(\theta - \phi)} + z, \quad (2)$$

where (u, v) defines an imager fixed coordinate system which rotates with the gantry. The imager center is the origin of the uv plane. The focal distance (F) is the SID. Details on how to derive Eqs. (1) and (2) are summarized in Appendix A. Considering the variations ($\Delta\phi$, ΔR , ΔF , Δu , and Δv) of the geometric parameters (ϕ , R , F , u , and v), the projections become

$$u + \Delta u = \frac{-(F + \Delta F)r \cos[\theta - (\phi + \Delta\phi)]}{(R + \Delta R) + r \sin[\theta - (\phi + \Delta\phi)]}, \quad (3)$$

$$v + \Delta v = \frac{(F + \Delta F)(z_B - z)}{(R + \Delta R) + r \sin[\theta - (\phi + \Delta\phi)]} + z. \quad (4)$$

Assuming that the imager might be tilted along an imaginary axis p , which is at an angle α with u axis, the imaginary axes are defined as p and q , respectively, on the uv plane as shown in Fig. 3(a). The transformation between coordinates (p, q) and (u, v) is

$$\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}. \quad (5)$$

Further taking the tilted angle β into consideration [Fig. 3(b)], a new axis q' is defined accordingly with

$$q' = \left(1 + \frac{q}{F}\beta + \frac{1}{2}\beta^2\right)q. \quad (6)$$

In obtaining the above equation, it was assumed that the tilting angle is small ($<15^\circ$). Please refer to Appendix B for the derivation of Eq. (6).

It follows that the new projected coordinates (u_I, v_I) of the fiducial on a tilted imager are

$$\begin{aligned} \begin{pmatrix} u_I \\ v_I \end{pmatrix} &= \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} p \\ q' \end{pmatrix} \\ &= \begin{pmatrix} u \\ v \end{pmatrix} + \left(\frac{q}{F} + \frac{\beta}{2}\right)q\beta \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}. \end{aligned} \quad (7)$$

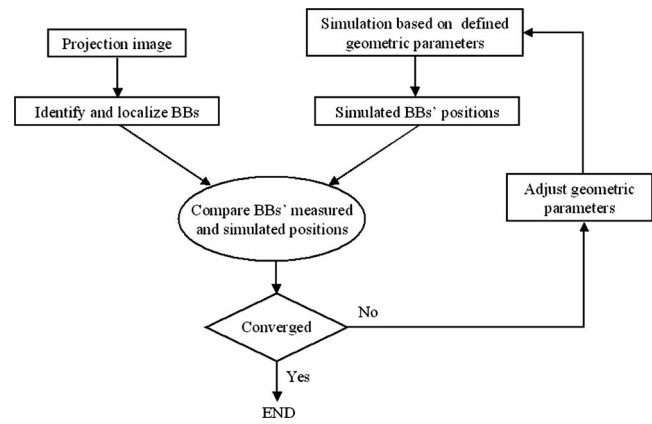


FIG. 4. Flow chart showing the algorithm of the analysis software.

II.C. Automated analysis software

From a projected image of the fiducials at a given gantry angle, eight geometric parameters could be quantitatively evaluated. These are: (1) Three parameters for the x-ray source position, namely the deviation of the gantry angle ($\Delta\phi$), the deviation of rotation radius (ΔR), and the deviation from the rotation plane (Δz); (2) three parameters for the position of the imager center, namely the deviation of the focal distance (ΔF), and the two translational offsets (Δu and Δv) in the uv plane; (3) two angular parameters (α and β) for the imager orientation.

An analysis software in C language was developed to estimate the eight parameters at any given gantry angle by comparing the measured and simulated BB locations in an iterative manner. The newly adjusted geometric parameters for the simulated BB location served as the inputs for the iteration (Fig. 4). The analysis software first identified and measured the BB locations (\bar{u}, \bar{v}) from the projection image in the presence of random background noise by intensity weighted averaging using

$$\begin{cases} \bar{u} = \frac{\sum_i u_i I_i}{\sum_i I_i} \\ \bar{v} = \frac{\sum_i v_i I_i}{\sum_i I_i} \end{cases}, \quad (8)$$

where I_i was the signal intensity of the pixel number i .²⁰ The software then determined the geometric parameters by minimizing the difference between the BBs measured and simulated locations from the simulation results based on Eqs. (3)–(7). For this purpose, the distance between the positions of the measured BB and its corresponding simulated BB was computed. The summation of these distances for all the BBs was used as the converging criterion for the optimization process. A hybrid optimization algorithm was adopted for this eight-parameter nonlinear optimization problem. Exhaustive search was applied first and possible solutions were compared in order to find the best 1000 seeds of parameter combinations for the simulated annealing algorithm to start with. Each step of the simulated annealing algorithm replaced the current solution by a random “nearby” solution

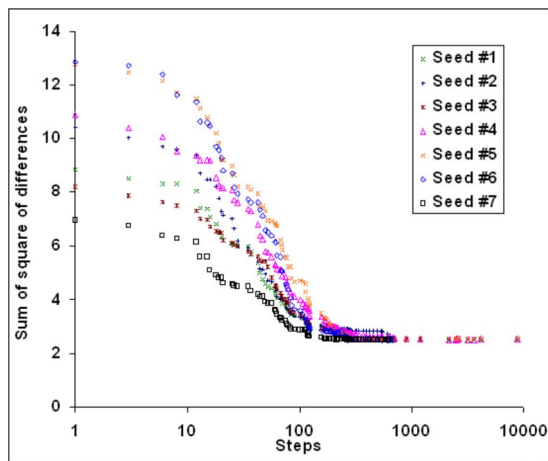


FIG. 5. Typical optimization converging curves.

chosen with a probability that depends on the difference between the corresponding function value and a global parameter (called the annealing temperature).^{21–24} The annealing temperature gradually decreased during the process. Finally the progress converged to the best estimate of the geometric parameters. Figure 5 illustrates several converging curves during the simulated annealing processes in analyzing a typical projection. This tool took about half a minute to analyze one projection image on a personal computer (DELL Precision 470 Workstation, 3.4 GHz CPU and 4 GB RAM). Typically, the sum of square distance differences converged to about 2.5 mm^2 . The average discrepancy for one fiducial was about 0.44 mm ($=\sqrt{2.5/13}$), which was around the size of an imager pixel.

II.D. Experimental evaluation

The developed *gQA* phantom and the automated analysis software (collectively called the *gQA* tool) were tested on a Varian Trilogy linac (Varian Medical Systems, Palo Alto, CA). The Trilogy is equipped with two image acquisition systems (IAS3): (1) MV EPID and (2) kV on-board imaging system (OBI). Both systems have a flat-panel detector with a matrix dimension of 1024×768 . The physical pixel sizes of the MV and kV imagers are 0.392 and 0.388 mm, respectively. The Trilogy is routinely maintained, and all the operating parameters are monitored and assured through a comprehensive QA program as suggested by TG 40.²⁵ The center of the *gQA* phantom was positioned at the nominal treatment isocenter as indicated by the room lasers. The relationship between the intersection of the room lasers and the mechanic isocenter of the linac was assumed to be maintained and calibrated in accordance with the routine QA practice.²⁵

II.D.1. Reproducibility of the analysis results

Projection images were acquired for MV EPID at a gantry angle of 0° (Varian convention) and at a SID of 1500 mm, the values were nominal. To ensure the *gQA* tool functions

properly in the worst-case scenario, the MV EPID was chosen here because the image quality of the MV acquisition was relatively inferior compared to that of the kV OBI. A 6 MV beam and 25 monitor units were used. The field size was $26 \times 20 \text{ cm}^2$. Twenty-one projections, 1 min apart, were made with the same imaging geometry and fed into the analysis software tool to estimate the geometric parameters of the imaging system. Although the 21 projection images were acquired under the same condition, they were associated with different random background noise. The purpose of this evaluation was to assess the reproducibility of the analysis results under different noisy environments.

II.D.2. Minimum detectable change in gantry angle

With the same setup, four projection images at nominal gantry angles of 0° , 0.1° , 0.2° , and 0.5° were taken. The projection images at nominal gantry angles of 0.1° , 0.2° , and 0.5° were subtracted from the one of 0° gantry angle to show the difference between the projection images made at slightly different angles. The purpose of this evaluation was to demonstrate the minimum change in gantry angle that could be detected from the projection image of the BBs.

II.D.3. Gantry angle offset

Further projection images were taken at nominal gantry angles of 0° , $\pm 0.1^\circ$, $\pm 0.2^\circ$, $\pm 0.3^\circ$, $\pm 0.5^\circ$, and $\pm 1.0^\circ$. This narrow range of gantry angle was chosen because an offset would affect a small gantry angle more than a large one. These images were then analyzed. The discrepancies between the nominal and estimated gantry angles were then compared in order to determine the gantry angle offset.

II.D.4. SID offset

Projection images were further acquired for MV EPID at a gantry angle of 180° but at different nominal SID of 1300, 1400, 1500, 1550, and 1600 mm. The best estimate of the SID was found by the analysis software. The discrepancies between the nominal and estimated SID were then compared in order to determine the SID offset.

II.D.5. Imager center offset

With the SID reset to 1500 mm, projection images were acquired with the imager center at (0, 0) and offset to (20, 20), (20, -20), (-20, -20), and (-20, 20) mm in the *uv* plane. The best estimate of the detector shift was found by the analysis software. The discrepancies between the actual and estimated offsets with and without systematic drift correction were then compared in order to determine the imager center offset.

II.D.6. Phantom positional shift

With the imager center reset to the origin, the *gQA* phantom was displaced -5.0 mm in each of the *x*, *y*, and *z* directions simultaneously. This was to mimic a condition where there were spatial drifts in the imaging geometry ($\Delta R, \Delta z, \Delta F, \Delta u, \Delta v$) in 3D space. When the phantom posi-

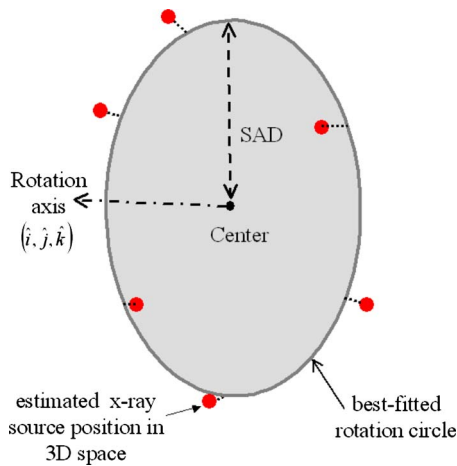


FIG. 6. Schematic diagram showing a best fitted rotation circle that has the shortest distances (the least sum distance square) from all the estimated x-ray sources in 3D space. The center of this circle is taken as the isocenter of the gantry rotation.

tion was shifted, the whole imaging system, including the source and imager position, was shifted in the opposite direction from the phantom's point of view. Projection images were then acquired at a SID of 1500 mm with gantry angles at 0° , 90° , 180° , and 270° . The same set of projection images was repeated with the phantom displaced 5.0 mm in the x, y, z direction and 10.0 mm in the same direction as well. The geometric parameters corresponding to these changes were estimated by the analysis software. The purpose of this evaluation was to demonstrate the accuracy of the *gQA* tool in detecting the variations in the imaging geometry.

II.D.7. Coincidence of the isocenters for the MV and kV beams

The center of the *gQA* phantom was repositioned at the nominal treatment isocenter as indicated by the room lasers. Projection images were acquired for MV EPID at every 10° of gantry angle for 360° and at a SID of 1500 mm resulting in 36 projection images in total. Projection images were also acquired for the OBI at every 0.56° of gantry angle for 360° and at a SID of 1500 mm resulting in 640 projection images in total (i.e. CBCT acquisition mode). The exposure factors were 125 kVp, 2 mAs, and the field size was $26 \times 20 \text{ cm}^2$ for the OBI projection image. The eight geometric parameters for each series of projected images were estimated and the

variation of each parameter for the full gantry rotation was examined. For each series of data, a best-fitted rotation circle was found such that it had the shortest distances (the least sum distance square) from all estimated x-ray sources in 3D space. The direction of the rotation circle was denoted by the unit vectors $(\hat{i}, \hat{j}, \hat{k})$, and its radius (SAD) was calculated. The rotation circle was taken as the rotation trajectory of the imaging system and its center as the isocenter (Fig. 6). The isocenters for the MV and kV beams were then found and compared.

II.D.8. Relationship among various geometric parameters

A change in one geometric parameter, for instance, the gantry angle, might affect other parameters of the imaging system as the whole imaging assembly is mounted on a laden gantry and robotic arms. There might be a subtle relationship between these parameters depending on the actual configuration and position of the imaging system. We studied this issue by scrutinizing the data obtained in Sec. II D 7 in which the projection images were acquired at different gantry angles. The aim was to see the changes in other parameters brought by the change in the gantry angle. We limited the studied range to be in $\pm 45^\circ$ because, beyond this range, the sagging of the gantry might overwhelm the subtle relationship amongst the different parameters. Furthermore, we studied the changes in parameters brought about by the specific parameter we varied in other evaluations such as in SID offset and imager off center (Secs. II D 4 and II D 5). All the results were compiled in a table to show the changes.

III. RESULTS

III.A. Reproducibility of the analysis results

Table I shows the variations of each of the eight geometric parameters of the MV imaging system for the 21 projection images with different background noise. The small standard deviation of each estimated geometric parameter indicated a good agreement of the data and the analysis results were highly reproducible. The imager center offset ($\Delta u, \Delta v$) was up to $(-0.63 \text{ mm}, 1.67 \text{ mm})$, which was due to a systematic drift of the MV imager.

TABLE I. The variations of each of the eight geometric parameters of the imaging system for the 21 projection images with different background noise. $\Delta\phi$, ΔR , and Δz are the deviations of the gantry angle, rotation radius and deviation from the rotation plane, respectively. ΔF , Δu , and Δv are the deviations of the source-to-detector distance, and the two translational offsets of the detector in the uv plane, respectively; α and β describe the orientation of the imager.

	$\Delta\phi$ ($^\circ$)	ΔR (mm)	Δz (mm)	ΔF (mm)	Δu (mm)	Δv (mm)	α ($^\circ$)	β ($^\circ$)
Min.	-0.39	-0.69	0.03	-0.48	-0.65	1.63	-0.28	-1.00
Max.	-0.35	-0.52	0.10	-0.24	-0.61	1.71	-0.11	-0.92
Mean	-0.37	-0.59	0.06	-0.37	-0.63	1.67	-0.22	-0.98
SD	0.01	0.04	0.02	0.06	0.01	0.02	0.05	0.02

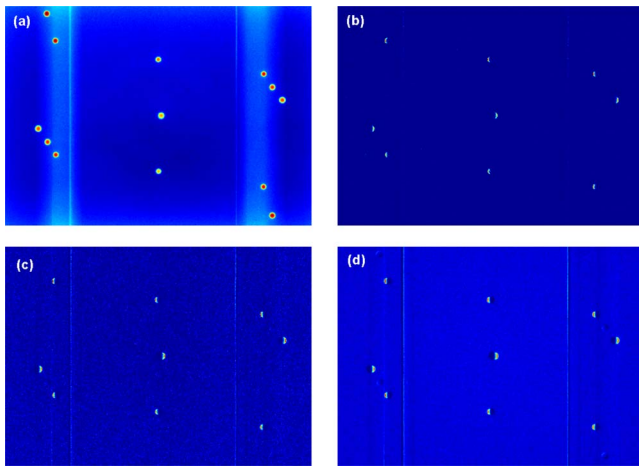


FIG. 7. (a) The projection image of the *gQA* phantom at gantry angle 0° . (b)–(d) show the image subtractions of the projection images at gantry angles 0.1° , 0.2° , and 0.5° from that at gantry angle 0° , respectively.

III.B. Minimum detectable change in gantry angle

Figure 7(a) shows the projection image at gantry angle 0° . Figures 7(b)–7(d) show the image subtractions of the projection images at the gantry angles 0.1° , 0.2° , and 0.5° from that at 0° , respectively. From the subtraction images, 9 out of 13 BBs' locations showed residual values indicating that 4 BBs were not differentiated by the projections made with small gantry angle changes and were totally subtracted out. In other words, there still remained 9 BBs that were “sensitive” enough to show the gantry change down to 0.1° .

III.C. Gantry angle offset

Figure 8 depicts the correlation of the nominal and estimated gantry angles in the range of $\pm 1^\circ$. The coefficient of determination (R^2) of 1.00 shows a very good correlation

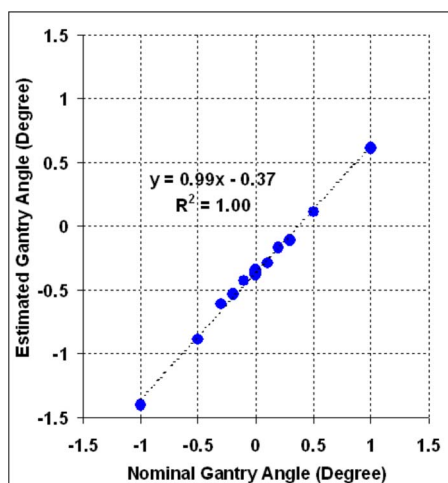


FIG. 8. The correlation of the nominal and estimated gantry angles in the range of $\pm 1^\circ$.

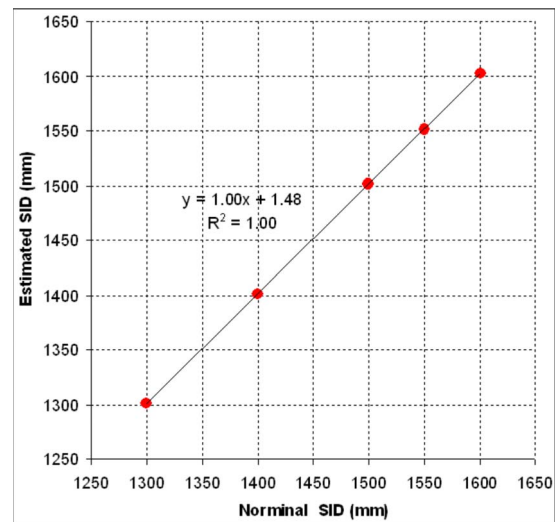


FIG. 9. The correlation of the nominal and estimated source-to-imager distance (SID) at a gantry angle of 180° .

between the two angles. From the regression line, a residual angle offset of 0.37° is evident for the nominal value of the gantry angle in the range studied.

III.D. SID offset

Figure 9 shows the correlation of the nominal and estimated SID in the range of 1300–1600 mm. The R^2 of 1.00 shows an almost perfect linearity between nominal and estimated SID. The regression line indicates that there is a consistent offset of 1.48 mm to the nominal SID in the range studied.

III.E. Imager center offset

Figure 10 shows the actual and estimated offset positions of the MV imager center on the (u, v) imager coordinate plane. A systematic drift of the detector was noted as observed in the previous evaluation II.D.1. If this systematic drift was corrected, the net estimated offset positions of the imager were close to the actual offsets within ± 1 mm.

III.F. Phantom positional shift

Table II lists the estimated relative shifts of the imaging geometry for the displaced phantoms with different magnitudes of displacement at the four principal orthogonal gantry angles. Depending on the gantry angle, ΔR , Δz , ΔF , Δu , and Δv reflected the displacement of the phantom in each of the x, y, z directions. The signs of the variations depend on the relative position between the phantom and the imaging system governed by the gantry angle. The maximum discrepancy between the actual and estimated shift was less than 1.6 mm ($1.6 = |10 - 8.4|$), which came from the differences between phantom shift of 10 mm and the estimated ΔR .

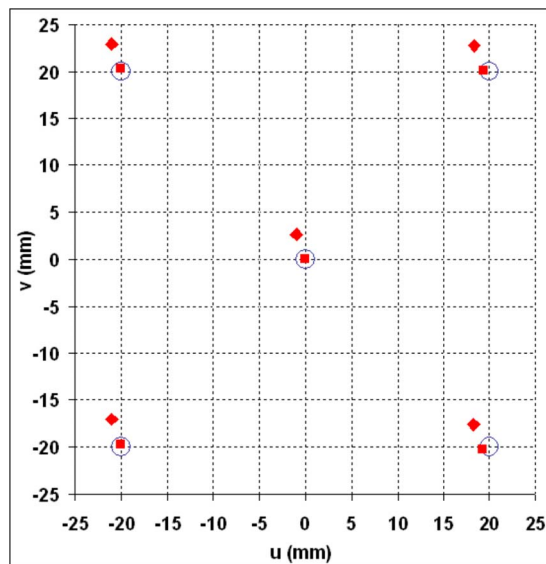


FIG. 10. The estimated shift (\blacklozenge) and estimated shift after the correction of the systematic drift (\blacksquare) compared to the actual offset of imager center (\circ) on the (u, v) imager coordinate plane.

III.G. Coincidence of the isocenters for the MV and kV beams

Table III lists the variation of each of the eight geometric parameters estimated for the MV EPID and kV OBI for a complete gantry rotation. The standard deviation of each estimated geometric parameter for the MV EPID and kV OBI were small indicating a good agreement of the same geometric parameter at different gantry angles. The orientations of the MV and kV rotation plane were found to be (0.00010, 0.00006, 1.00000) and (0.00030, 0.00001, 1.00000), respectively. The estimated isocenters for the MV and kV beams were at $(-0.95, -0.84, -0.25)$ mm and $(-0.97, -1.21, -0.13)$ mm in the reference coordinate. The coincidence of the two isocenters was well within 0.5 mm.

III.H. Relationship among various geometric parameters

Table IV lists the changes of geometric parameters of the imaging system brought about by the specific parameters varied in different evaluations. A change in one specific parameter could, to a certain extent, cause changes in other parameters depending on the actual configuration or the position of the imaging system, however, these changes were small and assumed not to affect the results of the experimental evaluations.

IV. DISCUSSION

We have designed a geometric QA phantom tailoring to an automated analysis to successfully estimate the geometric parameters of an on-board imaging system with a reasonable accuracy. The size of the phantom and the helical arrangement of the fiducials on the surface ensure the projected positions of the fiducials are well spaced on a detector area at any gantry angle for easy detection and identification. In principle, the more the number of the fiducials is, the more information that can be extracted for estimating the geometric parameters, but there will be higher chances of superposition among the fiducial projections. There is a trade-off between the number of fiducials and the minimum distance between the fiducial projections. Based on the result from the evaluation II.D.2, we concluded that 13 BBs were adequate for the present application and this *gQA* tool can detect an angular uncertainty of 0.1° .

Concerning the analysis software, the direct relationship between the 2D projected position of a fiducial and the idealized geometric parameters of an imaging system was used in an optimal manner for determining the geometric parameters. Theoretically, eight equations (from four fiducials' 2D projections) might be enough to solve this type of eight-parameter problem in a perfect mathematical model. In reality, the system is not perfect and several imperfect factors might occur: (1) The x-ray source is not a real geometry

TABLE II. The estimated relative drifts of the imaging geometry for the displaced phantoms with different magnitudes of displacement at four principal orthogonal gantry angles. ΔR , Δz , and ΔF are the deviation of the rotation radius, deviation from the rotation plane and deviations of the source-to-detector distance, respectively. Δu and Δv are the two translational offsets of the detector in the uv plane, respectively.

Phantom shift (mm)	Gantry angle ($^\circ$)	ΔR (mm)	Δz (mm)	ΔF (mm)	Δu (mm)	Δv (mm)
$\Delta x = -5.0$	0	3.7	-5.4	-0.2	4.6	5.3
$\Delta y = -5.0$	90	4.5	-5.1	0.0	-3.9	5.4
$\Delta z = -5.0$	180	-4.2	-5.4	-0.2	-4.5	5.3
	270	-4.6	-5.6	0.1	4.0	5.2
$\Delta x = 5.0$	0	-5.6	4.6	0.0	-4.6	-4.8
$\Delta y = 5.0$	90	-4.5	4.8	0.0	5.6	-4.8
$\Delta z = 5.0$	180	5.4	4.8	-0.2	4.7	-4.8
	270	4.9	4.7	0.2	-5.5	-4.8
$\Delta x = 10.0$	0	-8.4	9.7	0.0	-9.7	-9.9
$\Delta y = 10.0$	90	-9.2	10.2	0.2	8.6	-9.8
$\Delta z = 10.0$	180	8.4	10.4	-0.3	9.5	-9.6
	270	9.9	8.9	-0.4	-8.4	-10.4

TABLE III. The variations of each of the eight geometric parameters of the MV/kV imaging systems from projections acquired through a complete gantry rotation. $\Delta\phi$, ΔR , and Δz are the deviations of the gantry angle, rotation radius and deviation from the rotation plane, respectively. ΔF , Δu , and Δv are the deviations of the focal distance, and the two translational offsets of the detector in the uv plane, respectively.

		$\Delta\phi$ (°)	ΔR (mm)	Δz (mm)	ΔF (mm)	Δu (mm)	Δv (mm)	α (°)	β (°)
MV	Min.	-0.4	-1.2	-0.8	-2.3	-1.0	1.3	-0.3	-1.4
	Max.	0.0	2.7	0.1	2.5	-0.5	2.2	0.3	-0.6
	Mean	-0.2	0.5	-0.3	0.2	-0.8	1.7	0.0	-1.0
	SD	0.1	1.4	0.3	1.8	0.2	0.3	0.2	0.2
kV	Min.	-0.3	0.2	-0.8	-2.2	-1.2	1.3	-0.3	-2.4
	Max.	0.1	4.9	0.2	2.3	0.9	2.5	0.4	-0.7
	Mean	-0.1	2.2	-0.2	-0.1	-0.1	1.9	0.0	-1.3
	SD	0.1	1.5	0.3	1.6	0.7	0.4	0.2	0.5

point and has a finite size, which leads to geometric penumbra; (2) the accuracy of the spatial location of the fiducial is limited by measurement; and (3) the imager pixel has a finite size and is associated with noise. These imperfections render us to summon more equations in solving the problem. This proposed analysis tool takes all fiducial projections into account in the optimization. In addition, Eqs. (3)–(7) indicate that this is a nonlinear problem and multiple local minima might possibly coexist with the global minimum. While it is not clear whether the estimation results from some image-based optimization approaches¹⁹ would be trapped in the local minimum, we adopted to use a global optimization algorithm of simulated annealing for our software to ensure a global minimum is always achievable. Moreover, the exhaustive search at the beginning of the process warrants well-distributed seeds for the simulated annealing and, in turn, saves optimization time significantly.

We have evaluated the *gQA* tool by estimating the geometric parameters of the imaging system under various imaging geometries. The results were highly reproducible and showed that the developed tool was responsive to all the changes introduced in the evaluation such as SID, detector offset, and phantom shift. The best estimates of these spatial parameters in different situations were well within 1.6 mm.

The estimation of the rotation plane and the isocenter by a series of projection images of the *gQA* phantom through a full gantry rotation is extremely useful in monitoring the geometric parameters pertinent to the flat-panel based CBCT and the coincidence of the MV and kV beams. We found that the coincidence of the MV/kV beam isocenters were 0.5 mm

comparable to the value of 0.8 mm which was obtained by the routine OBI QA exercise on the Trilogy. Comparing to the existing OBI QA exercise, our tool is more comprehensive giving a full picture of the geometric status and details on the rotation trajectories of both the x-ray sources (MV/kV) and their orientation of the rotation planes. Yoo *et al.*¹⁷ pointed out that their recommendation did not include a separate geometric QA measurement for the OBI CBCT, the developed tool would be a good choice to fill it this gap because of its simplicity, easy implementation and full automation of the analysis.

Just like any other phantom-based QA studies,^{13–15,19} one limitation of the present work lies on the mechanical imprecision in the positions of the BBs since the simulation of the BBs position assumes a perfect alignment of the BBs in the designated skewed helical trajectory. The BBs in the present *gQA* phantom are accurate to within 0.5 mm in their designated position and is considered adequate for the current application.

A simple guideline is recommended here on how to use this *gQA* tool in a clinic. First, set up the phantom to the room lasers, which is used to provide a reference to the room coordinates. A level can be used or even embedded in the phantom to assure its geometric setup. Second, kV and/or MV projection images are acquired at a designated angle. Each (kV or MV) imaging system is described by the eight geometric parameters. To obtain the eight parameters of the imaging system (kV or MV) at a designated angle, in principle, a single projection measurement at that angle is sufficient. Last, the projection image is fed to the analysis soft-

TABLE IV. The changes of geometric parameters of the imaging system brought about by the specific parameters varied in different evaluations. $\Delta\phi$, ΔR , and Δz are the changes in the gantry angle, rotation radius and distance from the rotation plane, respectively. ΔF , Δu , and Δv are the changes of the focal distance, and the two translational offsets of the imager in the uv plane, respectively; α and β describe the changes in the tilting of the imager.

Evaluation	Parameter varied	Range varied	$\Delta\phi$ (°)	ΔR (mm)	Δz (mm)	ΔF (mm)	Δu (mm)	Δv (mm)	α (°)	β (°)
II.D.4	SID	1300–1600 mm	0.01	1.06	0.08	...	0.23	0.75	0.34	0.13
II.D.5	Imager off center	± 20 mm	0.00	0.27	0.08	0.12	0.61	0.25
II.D.7	Gantry angle	$\pm 45^\circ$...	0.28	0.05	0.50	0.02	0.06	0.13	0.05

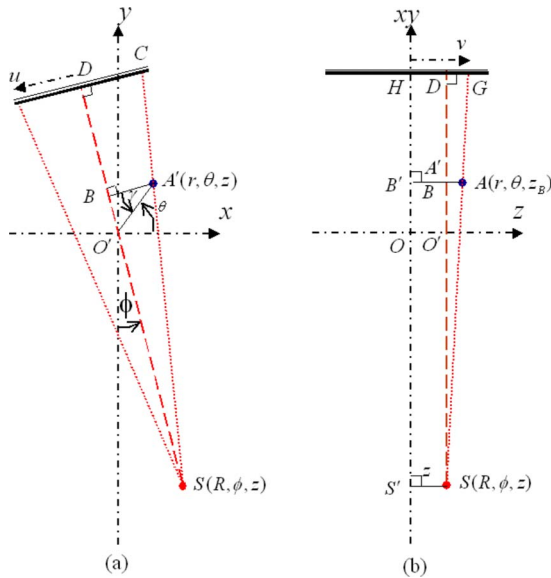


FIG. 11. Geometric relationship between a fiducial's 3D position and its projection on the imager.

ware to calculate the geometric parameters. A more comprehensive approach would be acquiring the projections during a continuous gantry rotation (kV CBCT/MV arc mode delivery). In that case, the geometric parameters at different gantry angles could be evaluated. Furthermore, the isocenters of the imaging system could also be calculated by the software if three or more projections are available.

V. CONCLUSIONS

A new geometric QA phantom and an automated analysis software have been developed to estimate the geometric status of a MV or kV on-board imager. This provides us a simple, robust and objective way to probe and monitor the geometric status of an imaging system in a fully automatic process, and facilitate routine QA workflow in a clinic.

ACKNOWLEDGMENTS

This work was supported in part by grants from the Department of Defense (PC040282), the National Cancer Institute (1R01 CA104205), and the Komen Breast Cancer Foundation (BCTR0504071). Support from the Sir Robert Black Postdoctoral Fellowship and the Hong Kong Hospital Authority Overseas Training Allowance for the second author are also gratefully acknowledged.

APPENDIX A: DERIVATION OF THE FIDUCIAL LOCATION (U, V)

When the x-ray source $S(R, \phi, z)$ is at a gantry angle of ϕ , an arbitrary object $A(r, \theta, z_B)$ is projected to C on imager u axis centering at D . Figure 11(a) provides a 2D view on an xy plane where the source S locates (z). $A'(r, \theta, z)$ is the projection of A on this plane. Using the similar triangle relationship between the triangles $SA'B$ and SCD , we have Eq. (A1)

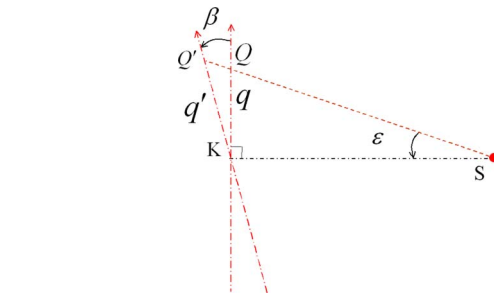


FIG. 12. Schematic diagram showing the tilting of the imager upon the p axis.

$$\frac{|CD|}{|A'B|} = \frac{|SD|}{|SB|}, \quad (\text{A1})$$

while $|A'B| = r \cos \gamma$ is the distance from A' to the projection central line (SO') and SD is the source-to-imager distance (F). SB is the distance from source to the projection on the central line. It is calculated by Eq. (A2)

$$|SB| = R + r \sin \gamma, \quad (\text{A2})$$

given

$$\begin{aligned} \gamma &= \angle BA'O' = \angle A'O'S - \angle A'BO' \\ &= \left[\left(\theta + \frac{\pi}{2} \right) - \phi \right] - \frac{\pi}{2} = \theta - \phi. \end{aligned} \quad (\text{A3})$$

So we have

$$|CD| = F \frac{|A'B|}{R + r \sin \gamma} = \frac{Fr \cos(\theta - \phi)}{R + r \sin(\theta - \phi)}$$

or

$$u = \frac{-Fr \cos(\theta - \phi)}{R + r \sin(\theta - \phi)}.$$

The negative sign was added due to the definition of u axis.

A side view perpendicular to the beam axis is shown in Fig. 11(b). Now A' and B are overlapped on this view. Based on the similarity between triangles SAB and SGD , the magnification factor is

$$\frac{|DG|}{|AB|} = \frac{|SD|}{|SB|}$$

while $|AB| = |AB'| - |SS'| = z_B - z$, $|SD| = F$, $|SB| = R + r \sin \gamma$.

So the projection on the v axis is

$$v = |GH| = |DG| + z = \frac{F(z_B - z)}{R + r \sin(\theta - \phi)} + z.$$

APPENDIX B: DERIVATION OF THE TILTED AXIS q'

After a tilting of β , the projection at Q moved to Q' , and the coordinate on q axis becomes a coordinate on q' axis. As shown in Fig. 12, from the triangle KQQ' , we have

$$\frac{|KQ|}{\sin(\angle Q'QK)} = \frac{|KQ'|}{\sin(\angle QQ'K)}.$$

Because

$$\angle Q'QK = \pi - \left(\beta + \frac{\pi}{2} \right) - \varepsilon = \frac{\pi}{2} - (\beta + \varepsilon)$$

and

$$\angle QQ'K = \frac{\pi}{2} + \varepsilon,$$

so

$$\frac{q'}{q} = \frac{|KQ'|}{|KQ|} = \frac{\sin\left(\frac{\pi}{2} + \varepsilon\right)}{\sin\left(\frac{\pi}{2} - (\beta + \varepsilon)\right)} = \frac{\cos(\varepsilon)}{\cos(\beta + \varepsilon)}.$$

Use small angle approximation, $\cos(\varepsilon) \approx 1 - 1/2\varepsilon^2$ and $\cos(\beta + \varepsilon) \approx 1 - 1/2(\beta + \varepsilon)^2$,

$$\begin{aligned} \frac{q'}{q} &\approx \frac{1 - \frac{1}{2}\varepsilon^2}{1 - \frac{1}{2}(\beta + \varepsilon)^2} \approx \left[1 - \frac{1}{2}\varepsilon^2 \right] \left[1 + \frac{1}{2}(\beta + \varepsilon)^2 \right] \\ &\approx 1 + \beta\varepsilon + \frac{1}{2}\beta^2. \end{aligned}$$

Approximately, $\varepsilon \approx q/F$.

So

$$q' \approx \left(1 + \beta\varepsilon + \frac{1}{2}\beta^2 \right) q.$$

^{a)} Author to whom correspondence should be addressed. Electronic mail: lei@reyes.stanford.edu

¹ B. Sorcini and A. Tilikidis, "Clinical application of image-guided radiotherapy, IGRT (on the Varian OBI platform)," *Cancer Radiother* **10**, 252–257 (2006).

² C. A. McBain, A. M. Henry, J. Sykes, A. Amer, T. Marchant, C. M. Moore, J. Davies, J. Stratford, C. McCarthy, B. Porritt, P. Williams, V. S. Khoo, and P. Price, "X-ray volumetric imaging in image-guided radiotherapy: The new standard in on-treatment imaging," *Int. J. Radiat. Oncol., Biol., Phys.* **64**, 625–634 (2006).

³ L. Xing, B. Thorndyke, E. Schreibmann, Y. Yang, T. F. Li, G. Y. Kim, G. Luxton, and A. Koong, "Overview of image-guided radiation therapy," *Med. Dosim.* **31**, 91–112 (2006).

⁴ Y. Yang, E. Schreibmann, T. Li, C. Wang, and L. Xing, "Evaluation of on-board kV cone beam CT (CBCT)-based dose calculation," *Phys. Med. Biol.* **52**, 685–705 (2007).

⁵ D. Letourneau, J. W. Wong, M. Oldham, M. Gulam, L. Watt, D. A. Jaffray, J. H. Siewerdsen, and A. A. Martinez, "Cone-beam-CT guided radiation therapy: Technical implementation," *Radiother. Oncol.* **75**, 279–

286 (2005).

⁶ W. Mao, T. Li, N. Wink, and L. Xing, "CT image registration in sinogram space," *Med. Phys.* **34**, 3596–3602 (2007).

⁷ C. Huntzinger, P. Munro, S. Johnson, M. Miettinen, C. Zankowski, G. Ahlstrom, R. Glettig, R. Filliberti, W. Kaissl, M. Kamber, M. Amstutz, L. Bouchet, D. Klebanov, H. Mostafavi, and R. Stark, "Dynamic targeting image-guided radiotherapy," *Med. Dosim.* **31**, 113–125 (2006).

⁸ T. Li, L. Xing, P. Munro, C. McGuinness, M. Chao, Y. Yang, B. Loo, and A. Koong, "Four-dimensional cone-beam computed tomography using an on-board imager," *Med. Phys.* **33**, 3825–3833 (2006).

⁹ M. Oldham, D. Letourneau, L. Watt, G. Hugo, D. Yan, D. Lockman, L. H. Kim, P. Y. Chen, A. Martinez, and J. W. Wong, "Cone-beam-CT guided radiation therapy: A model for on-line application," *Radiother. Oncol.* **75**, 271–278 (2005).

¹⁰ D. A. Jaffray, J. H. Siewerdsen, J. W. Wong, and A. A. Martinez, "Flat-panel cone-beam computed tomography for image-guided radiation therapy," *Int. J. Radiat. Oncol., Biol., Phys.* **53**, 1337–1349 (2002).

¹¹ J. Lu, T. M. Guerrero, P. Munro, A. Jeung, P.-C. M. Chi, P. Balter, X. R. Zhu, R. Mohan, and T. Pan, "Four-dimensional cone beam CT with adaptive gantry rotation and adaptive data sampling," *Med. Phys.* **34**, 3520–3529 (2007).

¹² M. B. Sharpe, D. J. Moseley, T. G. Purdie, M. Islam, J. H. Siewerdsen, and D. A. Jaffray, "The stability of mechanical calibration for a kV cone beam computed tomography system integrated with linear accelerator," *Med. Phys.* **33**, 136–144 (2006).

¹³ B. E. H. Claus, presented at the *Proceedings of SPIE*, 2006 (unpublished).

¹⁴ Y. Cho, D. J. Moseley, J. H. Siewerdsen, and D. A. Jaffray, "Accurate technique for complete geometric calibration of cone-beam computed tomography systems," *Med. Phys.* **32**, 968–983 (2005).

¹⁵ A. Rougee, C. Picard, C. Ponchut, and Y. Troussset, "Geometrical calibration of x-ray imaging chains for three-dimensional reconstruction," *Comput. Med. Imaging Graph.* **17**, 295–300 (1993).

¹⁶ R. Fahrig and D. W. Holdsworth, "Three-dimensional computed tomographic reconstruction using a C-arm mounted XRII: Image-based correction of gantry motion nonidealities," *Med. Phys.* **27**, 30–38 (2000).

¹⁷ S. Yoo, G. Y. Kim, R. Hammoud, E. Elder, T. Pawlicki, H. Guan, T. Fox, G. Luxton, F. F. Yin, and P. Munro, "A quality assurance program for the on-board imagers," *Med. Phys.* **33**, 4431–4447 (2006).

¹⁸ K. Yang, A. L. Kwan, D. F. Miller, and J. M. Boone, "A geometric calibration method for cone beam CT systems," *Med. Phys.* **33**, 1695–1706 (2006).

¹⁹ M. Mamalui-Hunter, H. Li, and D. A. Low, presented at the *AAPM 49th Annual Meeting*, Minneapolis, MN, 2007 (unpublished).

²⁰ L. Dong, A. Shiu, S. Tung, and A. Boyer, "Verification of radiosurgery target point alignment with an electronic portal imaging device (EPID)," *Med. Phys.* **24**, 263–267 (1997).

²¹ S. Webb, "Optimization of conformal radiotherapy dose distributions by simulated annealing," *Phys. Med. Biol.* **34**, 1349–1370 (1989).

²² A. Pugachev and L. Xing, "Incorporating prior knowledge into beam orientation optimization in IMRT," *Int. J. Radiat. Oncol., Biol., Phys.* **54**, 1565–1574 (2002).

²³ I. I. Rosen, K. S. Lam, R. G. Lane, M. Langer, and S. M. Morrill, "Comparison of simulated annealing algorithms for conformal therapy treatment planning," *Int. J. Radiat. Oncol., Biol., Phys.* **33**, 1091–1099 (1995).

²⁴ J. G. Li, A. L. Boyer, and L. Xing, "Clinical implementation of wedge filter optimization in three-dimensional radiotherapy treatment planning," *Radiother. Oncol.* **53**, 257–264 (1999).

²⁵ G. J. Kutcher *et al.*, "Comprehensive QA for radiation oncology: Report of AAPM Radiation Therapy Committee Task Group 40," *Med. Phys.* **21**, 581–618 (1994).

Fast internal marker tracking algorithm for onboard MV and kV imaging systems

W. Mao, R. D. Wiersma, and L. Xing^{a)}

Department of Radiation Oncology, Stanford University School of Medicine, Stanford, California 94305-5847

(Received 12 October 2007; revised 31 January 2008; accepted for publication 26 February 2008; published 23 April 2008)

Intrafraction organ motion can limit the advantage of highly conformal dose techniques such as intensity modulated radiation therapy (IMRT) due to target position uncertainty. To ensure high accuracy in beam targeting, real-time knowledge of the target location is highly desired throughout the beam delivery process. This knowledge can be gained through imaging of internally implanted radio-opaque markers with fluoroscopic or electronic portal imaging devices (EPID). In the case of MV based images, marker detection can be problematic due to the significantly lower contrast between different materials in comparison to their kV-based counterparts. This work presents a fully automated algorithm capable of detecting implanted metallic markers in both kV and MV images with high consistency. Using prior CT information, the algorithm predefines the volumetric search space without manual region-of-interest (ROI) selection by the user. Depending on the template selected, both spherical and cylindrical markers can be detected. Multiple markers can be simultaneously tracked without indexing confusion. Phantom studies show detection success rates of 100% for both kV and MV image data. In addition, application of the algorithm to real patient image data results in successful detection of all implanted markers for MV images. Near real-time operational speeds of ~ 10 frames/sec for the detection of five markers in a 1024×768 image are accomplished using an ordinary PC workstation. © 2008 American Association of Physicists in Medicine. [DOI: [10.1118/1.2905225](https://doi.org/10.1118/1.2905225)]

Key words: fiducial tracking, IGRT, image guidance

I. INTRODUCTION

Modern conformal radiation therapy techniques, such as intensity-modulated radiation therapy (IMRT), can provide radiation doses that closely conform to the tumor dimensions while sparing sensitive structures.^{1,2} To be optimally effective, these techniques require a high geometric precision in both tumor localization and patient treatment setup. The presence of inter and intrafraction organ motion uncertainties can therefore reduce the benefit of using a highly conformal radiotherapy technique. For instance, intrafractional respiratory or prostate based tumor motion can lead to tumor displacements up to 2–3 cm over the course of routine radiotherapy.^{3–9} The use of image-guided radiation therapy (IGRT) is a promising candidate to ensure proper targeting in radiation treatment deliveries.¹⁰ Due to the dynamical nature of human anatomy, it is most advantageous when IGRT can be performed in real-time in order to ensure an accurate delivery of the planned conformal dose distribution.¹¹

Several methods of obtaining real-time tumor position are available, and these can be categorized as being either indirect (external surrogate-based) or direct (fiducial/image) in nature. In general, indirect tumor location methods, such as external skin marker tracking or breath monitoring techniques, rely on the correlation between external body parameters and the tumor.^{5,12} In reality, the relationship between external parameters and internal organ motion is complex

and a large uncertainty may be present in predicting the tumor location based on external signals. A direct tumor position measurement is therefore highly desirable for therapeutic guidance. In the last decade, a number of direct real-time 3D tumor tracking methods have been implemented, primarily using fluoroscopy^{5,11,13} or magnetic field localization.¹⁴ In addition, the feasibility of using an electronic portal imaging device (EPID) and stereoscopic x-ray imaging for tumor tracking has been explored.^{3,5,6,15–26}

To be clinically useful, an internal marker tracking algorithm should reliably segment markers from varying complex anatomic image backgrounds, be able to track multiple markers simultaneously without indexing confusion, and be able to operate at near real-time speeds. Generally, marker segmentation algorithms based on pixel intensity tend to fail when markers are in the vicinity of high contrast structures such as bone. A more reliable solution is the use of template matching, as demonstrated by Shirato *et al.*, in the tracking of a spherical gold marker using multiple kV fluoroscopic imaging systems.²⁵ Tang *et al.* further extended the template matching technique to include detection of cylindrically shaped markers on an in-house built stereoscopic kV imaging system.²⁷ In their algorithm, the user first manually located and defined a region of interest (ROI) around each marker and determined the orientation of every cylindrical marker, and then a template matching algorithm was applied to the ROIs. The normalized cross correlation between the

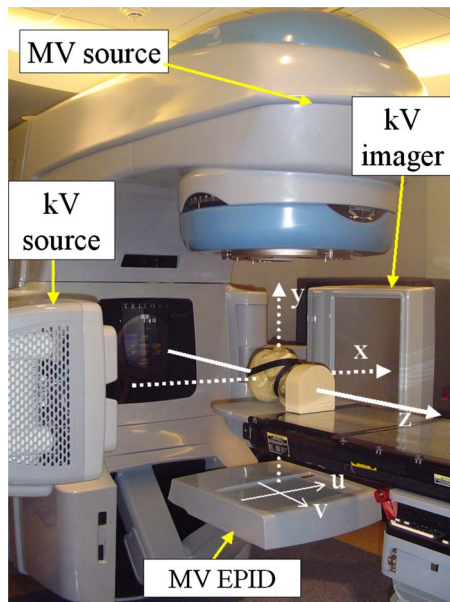


FIG. 1. Varian Trilogy with kV and MV imagers in extended positions. The system's frame of reference is denoted by arrows. A head phantom is located on the couch.

template and correlate pixels was calculated for every pixel and the highest cross correlation yielded the marker location. The detection failure rate for their method could be up to 12%.

As of yet, few works have presented algorithms suitable for tracking internal markers reliably using MV image data.¹⁸ This is a consequence of the low inherent contrast between different materials in MV based images, making image based marker detection difficult.

This work proposes a novel pattern matching algorithm specifically designed to work with both kV and MV imaging systems. Unlike previous algorithms reported by Shirato *et al.* and Tang *et al.* where only the cross correlation is used for marker detection, this algorithm employs a criteria system based not on only the correlation score, but also the scaling factor and their combination. This is found to allow 100% successful marker detection rates, even on low contrast MV images. This algorithm is highly desirable since it can combine MV imaging with kV fluoroscopy imaging to locate real-time 3D tumor position during the actual radiotherapy process, as recently demonstrated by Wiersma *et al.*²⁸ Unlike other tracking techniques, which require two or more kV sources for 3D marker positioning,^{3,5,6,25–27,29} this technique has the inherent benefit in that only one kV source is required for full 3D marker geometric information since the actual MV treatment beam is also used for positioning.

II. MATERIALS AND METHOD

II.A. Hardware setup

A Varian Trilogy (Varian Medical System, Palo Alto, CA) operating in the 6 MV photon mode was used for the study. Images of the MV beam were acquired by an aSi EPID

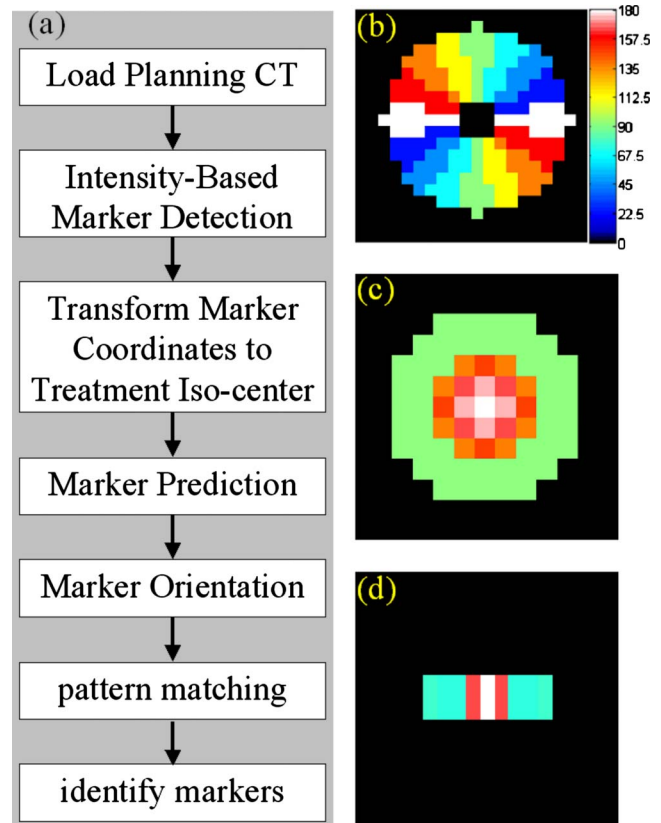


FIG. 2. (a) Flow chart algorithm computational path. (b) To the right is shown the eight-bin orientation segmentation filter, (c) spherical pattern, and (d) a cross section pattern for a vertical cylindrical marker.

(PortaVision aS-500, Varian Medical System, Palo Alto, CA) attached to the LINAC, as shown in Fig. 1. The kV imaging was accomplished using the onboard imaging system located perpendicular to the treatment beam (Fig. 1). The kV imaging system consisted of a 125 kV x-ray tube together with an aSi flat panel imager (PaxScan 4030CB, Varian Medical Systems, Salt Lake City, UT). Pixel sizes of the kV and MV detectors were 0.388 mm and 0.392 mm, respectively. Both detectors had a resolution of 1024×768 , corresponding to an effective area of detection of approximately $40 \text{ cm} \times 30 \text{ cm}$. The source-to-axis distances (SAD) were set to 100 cm and source-to-imager distances (SID) were set to 150 cm for both MV and kV systems (Fig. 1).

The markers used were either stainless steel ball bearings (BB) or gold (Au) cylinders. The BB diameters varied from 1.57 mm to 4 mm, whereas the Au cylinders were 1.2 mm in diameter and 5 mm in length (North West Medical Physics Equipment, MED-TEC Company, Orange City, IA). Efficiency of the algorithm in segmenting markers from complex anatomical image background was tested using a head phantom (Fig. 1). The cylindrical fiducials were internally embedded into the phantom, whereas BBs were mounted externally on the phantom.

II.B. CT based region-of-interest definition

Prior planning CT knowledge can provide valuable marker information that can be used to reduce the marker

search space. Reduction of search space in turn removes unnecessary image processing resulting in increased fiducial detection speeds. As displayed by the flow diagram in Fig. 2, a simple intensity based search for the markers was initially done on the planning CT. Due to the large CT numbers of the metallic markers relative to other anatomical structures, the markers were easily segmented from the image background. The displacement vector relating the CT isocenter to the treatment isocenter was then used to transform the marker CT coordinates to that of the treatment isocenter such that each marker was given a 3D position (x_B, y_B, z_B) relative to the machine isocenter. The expected projection location (u, v) of each marker on either the kV detector or the EPID can then be predicted by the following relationships

$$u = F \frac{\cos(\phi)x_B + \sin(\phi)y_B}{R - \sin(\phi)x_B + \cos(\phi)y_B}, \quad (1)$$

$$v = F \frac{z_B}{R - \sin(\phi)x_B + \cos(\phi)y_B}, \quad (2)$$

where ϕ was the gantry angle, R was SAD, and F was SID. As seen in Fig. 1, the x -axis was in the lateral direction of patient couch, the y -axis was in the anterior-posterior direction, the z -axis was the superior-inferior direction, and the origin was the LINAC's isocenter. The coordinates of the imagers were defined in the (u, v) plane in Fig. 1.

Having located the expected positions of the marker projections, the search region was reduced to a small circular ROI around each marker projection center. Typically, a ROI with a radius of 75 pixels on the imager, or about 2 cm around the fiducial position, was found to be adequate in locating the marker. For markers located near each other, overlapping ROIs would be grouped together. The fiducial search was then performed group by group in order to avoid redundancy.

II.C. Marker orientation

With a cylindrical marker of fixed length (l) and width (w), the unsymmetrical shape could lead to a host of different possible projection images depending on the marker's particular orientation relative to the source/imager setup. Parameters subject to variation were the marker's projection length and orientation (from 0° – 180°). The marker projection length might vary from “ $1.5 \times w$ ” (if the projection direction was along its longitudinal direction) to “ $1.5 \times l$ ” (if the projection direction was perpendicular to its longitudinal di-

rection) with an amplification factor of 1.5 (=SID/SAD). To take into account the different possible orientations, the 180° rotation was divided into a number of bins, as shown in Fig. 2(b). As seen in the figure, each of the eight bins represented two possible opposite directions, such that each bin covered an angle of 22.5° . The center of this orientation filter was placed on each pixel in a ROI group, and all adjacent pixels from the center pixel were then grouped into eight bins corresponding to their particular angle. The average intensity of each bin was computed and the bin with the highest number was taken as the orientation of the center pixel. It should be noted that the orientation results carried no information for most of the pixels and it was only valid for pixels on a cylindrical object. The eight-bin pattern was chosen based on the limited pixel resolution of a projected marker. Due to the small marker size, a detected projection was composed of a limited number of pixels and therefore mosaic in nature (typically with a width of ~ 5 pixels and a length of ~ 20 pixels or less). Therefore, for small angles of rotation, only a few pixels would change. Using bin numbers greater than eight generally did not lead to a more accurate angle determination since now the pixel noise could be the deciding factor in a particular angle bin.

II.D. Pattern matching algorithm

Different patterns were used for spherical and cylindrical markers. Due to rotational symmetry, a simple spherical marker pattern [Fig. 2(c)] was universally used for all BBs with the same physical diameter. For cylindrically shaped markers, a trapezoidal pattern [Fig. 2(d)] was used. The rationale behind this pattern was that even though a cylindrical marker projection might undergo a wide variety of rotational and length changes, the signal intensity distribution along the cross section of the cylinder projection remained constant. This cross section had a unique trapezoidlike pattern that was experimentally found to be dependent on the marker's width, but independent of projection length. Searching for these unique cross-sectional patterns along the orientation of every pixel in each ROI allowed for identification of fiducials.

Depending on the sought-after marker, either a spherical or cylindrical cross-sectional pattern $\{p_{i,j}\}$ is used for the matching procedure. At each pixel location (x, y) within a ROI group, a comparison was made between the pattern and the surrounding pixels ($\{f_{x+i,y+j}\}$). Two basic criteria were calculated: the square of the correlation coefficient $R_{x,y}^2$ and the scaling factor $H_{x,y}$

$$R_{x,y}^2 = \frac{\sum_{(i,j) \in \text{Pattern}} (f_{x+i,y+j} - \overline{f_{x,y}})(p_{i,j} - \overline{p})}{\sqrt{\sum_{(i,j) \in \text{Pattern}} (f_{x+i,y+j} - \overline{f_{x,y}})^2} \cdot \sqrt{\sum_{(i,j) \in \text{Pattern}} (p_{i,j} - \overline{p})^2}}, \quad (3)$$

$$H_{x,y} = \frac{\sum_{(i,j) \in \text{Pattern}} (f_{x+i,y+j} - \overline{f_{x,y}})(p_{i,j} - \overline{p})}{\sum_{(i,j) \in \text{Pattern}} (p_{i,j} - \overline{p})^2}, \quad (4)$$

where $\overline{f_{x,y}}$ was the average intensity of the pattern region around pixel (x,y) and \overline{p} was the average intensity of the pattern distribution as given by

$$\overline{f_{x,y}} = \frac{1}{N} \sum_{(i,j) \in \text{Pattern}} f_{x+i,y+j}, \quad (5)$$

$$\overline{p} = \frac{1}{N} \sum_{(i,j) \in \text{Pattern}} p_{i,j}, \quad (6)$$

with N being the total pixel number of the pattern. The square of the coefficient of correlation (R^2) for a linear regression could vary from 0 (no correlation) to 1 (perfect correlation). The scaling factor H indicated the relative intensity of the object compared to the background. As an example, in an ideal case, the image was scaled from the pattern after a background shift, $f_{x+i,y+j} = k \cdot p_{i,j} + b$, where k and b were constants. The pattern matching result would be $R_{xy}^2 = 1$ and $H_{x,y} = k$.

Because the cylindrical cross sectional pattern is unable to determine correct marker lengths, it was necessary to group adjacent qualified cross sections and reconstruct their overall lengths. If the calculated length was found to be longer than the maximum projection length, the feature would be rejected. Here the maximum projection length was defined by multiplying the actual physical marker length by the imager magnification factor ($=1.5$) plus a reasonable margin.

II.E. Marker identification

With multiple markers it was easy to confuse the individual marker labeling for kV or MV projections at different gantry angles. The simplest case was when only one marker exists in each ROI group. From Eqs. (1) and (2) there was a direct correlation of the ROI group to a specific marker. In the case of multiple markers existing in the same ROI group, the detected marker positions in this ROI group were correlated to the planning CT markers corresponding to this ROI group, while indexing was based on the shortest distances between detected and predicted marker projections by using Eqs. (1) and (2). At certain gantry angles, it was possible that two or more markers may be projected on the same (u,v) location. In this case the number of detected markers in the ROI group would be fewer than the number of expected markers, indicating projection overlapping. This was resolved by comparing the measured to the predicted projections. If one or more of the projections was found missing, but was calculated to be in close proximity to another marker, this projection would be double counted.

II.F. Experimental validation and patient data analysis

To evaluate the algorithm's efficiency when markers overlap or near different anatomical structures, a 360° gantry rotation was performed around a head phantom. Images were acquired every 0.56° for the kV imager (~ 640 images in total) and every 10° with the MV imager (36 images in total). Having obtained the projection locations for each marker, 3D spatial information could be calculated by using

projections at different gantry angles together with Eqs. (1) and (2). Particularly, the results from pairs of MV and kV projections were calculated.

As a preliminary test, this algorithm was also applied to actual patient images previously acquired for patient setup by onboard MV EPID. Five prostate patients with implanted cylindrical fiducial markers were treated in the past two years. Every patient had two or three fiducials (with a diameter of 1.2 mm and a length of 3 mm) implanted. They were treated on a conventional LINAC with MV EPID only. A total of 196 MV projection images were acquired for patient setup at anterior/posterior (AP) and lateral (LAT) directions before every treatment fraction. After all images were analyzed, the 3D spatial positions were calculated from every pair of AP and LAT images based on Eqs. (1) and (2).

An in-house software (C language) was specified to analyze projection images and obtain fiducial positions. All calculations were performed on a Dell Precision 470 workstation (3.4 GHz Xeon CPU and 4 GB RAM).

III. RESULTS

As a demonstration of the orientation and cross-sectional pattern matching, a MV image of a head phantom with five embedded cylindrical markers was examined [Fig. 3(a)]. For a selected marker enlarged in Fig. 3(b), the eight-bin orientations are shown in Fig. 3(c). As can be seen, for pixels located around the fiducial, the $\sim 45^\circ$ angle was favored, corresponding to the actual projection orientation. Having determined the orientation, a 90° rotation was made and the marker's cross section was segmented from the image back-

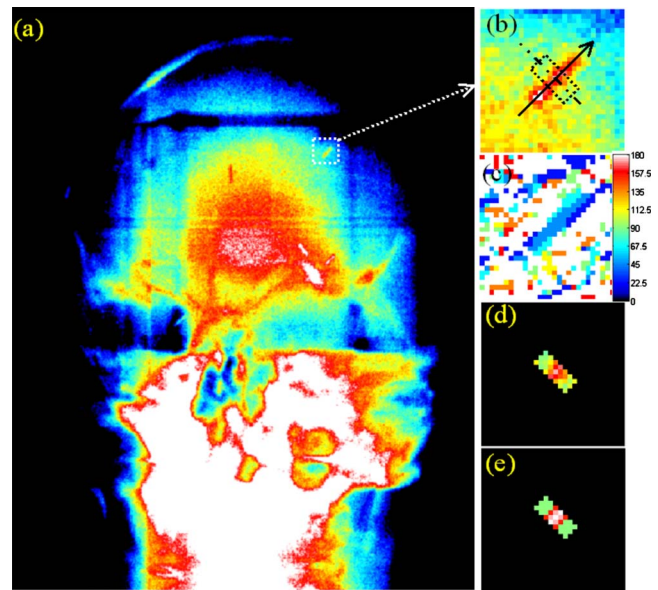


FIG. 3. (a) MV image of a head phantom with embedded Au cylindrical fiducial. (b) On the right shows magnified image of the selected marker displaying orientation and cross section (dotted line). (c) Orientation map with pixel intensity corresponding to adjacent intensity orientation for the selected marker. (d) Segmented cross-section. (e) Application of the cross-section pattern.

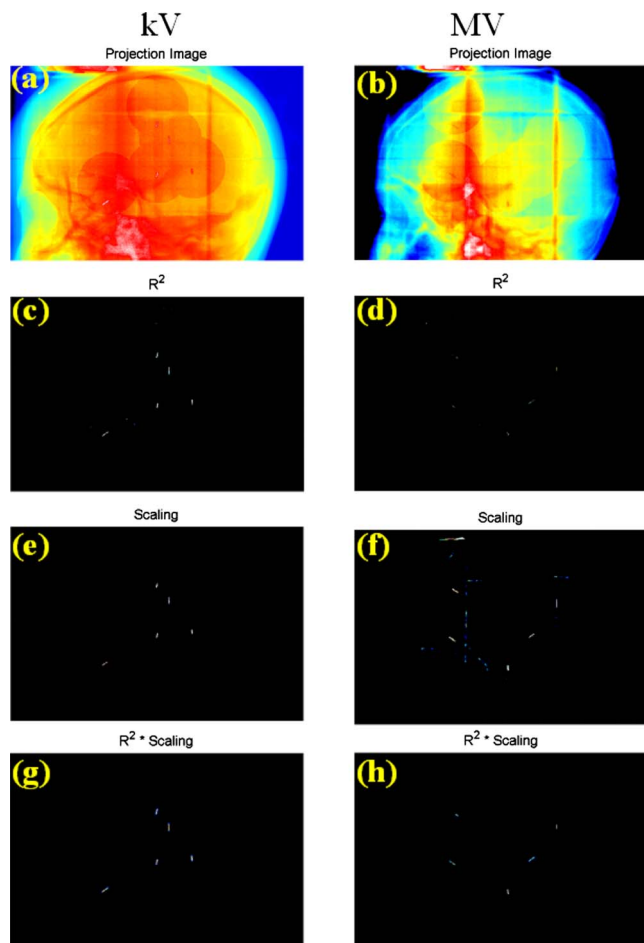


FIG. 4. Side-by-side kV (left column)/MV (right column) comparison of marker detection for five gold markers embedded in a head phantom. (a) and (b) Projection image and the predicted BB ROIs were highlighted, (c) and (d) R^2 results (threshold applied), (e) and (f) scaling factors (threshold applied), and (g) and (h) searching index = $R^2 \cdot$ scaling factor.

ground [Fig. 3(d)]. Then the cross-sectional pattern at 45° orientation [Fig. 3(e)] was applied to calculate $R^2_{x,y}$ and $H_{x,y}$.

Figure 4 is a side-by-side comparison of the kV and MV detection process for five cylindrical Au markers embedded in a head phantom. The kV and MV images were acquired at a MV gantry angle of 220° . It should be noted that the on-board kV x-ray source is always rotated 90° relative to the MV source (Fig. 1). Application of the algorithm uses prior planning CT information to first define the ROI for each fiducial, as shown by the circular highlighted regions in Figs. 4(a) and 4(b). After ROIs were defined and grouped, orien-

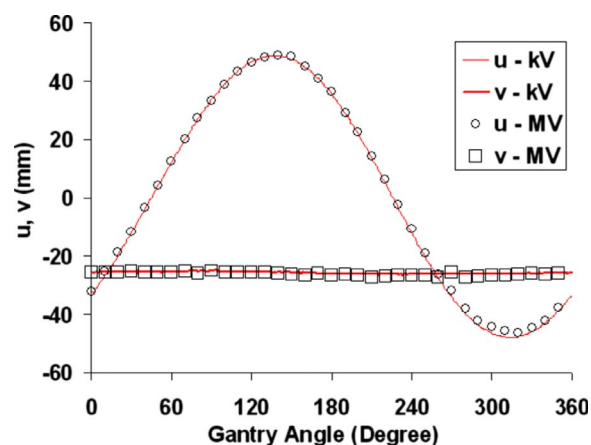


FIG. 5. Projection locations of one cylindrical fiducial on kV and MV images as functions of kV and MV gantry angles, respectively.

tation and pattern matching were applied to each ROI group and resulted in only image data that conformed to the cross-sectional pattern [Figs. 4(c)–4(f)]. The scaling of the ROIs led to greater background image segmentation for the kV case [Fig. 4(e)] over the MV case [Fig. 4(f)]. The product of the scaling factor H and the correlation R^2 led to complete segmentation of the markers, as displayed in Figs. 4(g) and 4(h). In this case a threshold of 0.6 and 0.006 for R^2 and H , respectively, were determined by previous trials and used through all our analyses.

Both kV and MV projections were analyzed using our detection algorithm, where it was found that all five markers were correctly detected for every kV and MV projection. Figure 5 plots the (u, v) coordinates on both the kV and MV detectors for one of the markers. Spatial positions of every marker were calculated for all MV projection images with their corresponding kV partner images gathered over the 360° gantry rotation. Table I summarizes these results, where it can be seen that the standard deviation of the location is better 0.5 mm and the errors are within 1 mm.

All patient images were analyzed similarly. In all cases the detection success rate was 100%. Figure 6 displays AP (left column) and LAT (right column) images of three markers embedded in a prostate. Three-dimensional positions of markers were calculated from pairs of AP and LAT projections. Figure 7 shows a fiducial marker's various 3D positions for 14 fractions. It should be noted that those images were taken at the beginning of patient setup and these positions were not yet the treatment position.

TABLE I. Variation of FM 3D positions calculated from MV/kV pair projections at 36 different gantry angles.

	FM #1			FM #2			FM #3			FM #4			FM #5		
	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z
Mean	-22.9	22.3	-17.2	20.7	24.1	0.0	2.5	-31.9	2.5	-17.3	-32.7	22.7	19.5	22.8	30.4
Max–Mean	0.7	0.7	0.3	0.4	0.6	0.4	0.5	0.6	0.3	0.6	0.7	0.3	0.5	0.6	0.6
Min–Mean	-0.6	-0.7	-0.3	-0.7	-0.6	-0.3	-1.0	-0.6	-0.3	-0.8	-0.8	-0.4	-0.7	-0.9	-0.4
Standard Dev	0.4	0.4	0.2	0.3	0.4	0.2	0.4	0.4	0.2	0.4	0.4	0.2	0.3	0.4	0.2

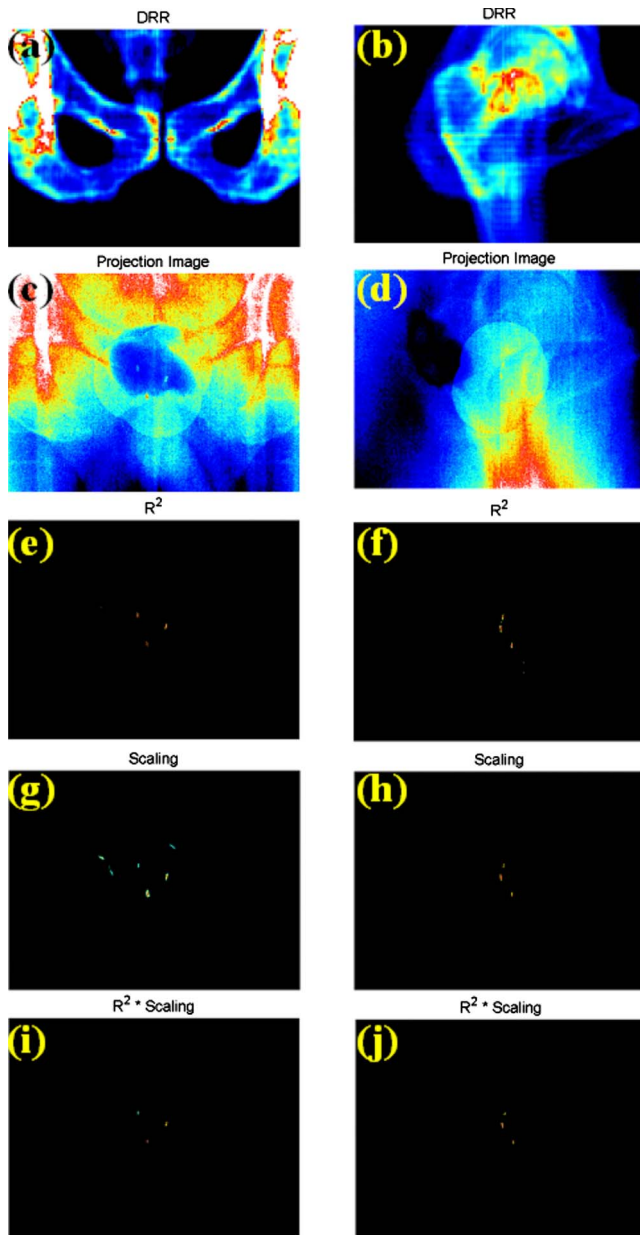


FIG. 6. Analyzing five cylindrical fiducials in a prostate patient MV image at AP (left column) and LAT (right column) directions. (a) and (b) DRR images, (c) and (d) Projection image and the predicted BB ROIs were highlighted, (e) and (f) R^2 results (threshold applied), (g) and (h) scaling factors (threshold applied), and (i) and (j) searching index= $R^2 \cdot$ scaling factor.

In the case for symmetrical markers such as BBs, the algorithm can be easily modified by using the spherical pattern, as shown in Fig. 2(c). This detection is more simplistic compared to cylindrical marker detection since the BB projection image is independent of the BB's orientation. More than 2000 combined kV and MV projection images were made of four spherical BB markers (Fig. 8). For all images the algorithm was able to successfully detect the markers.

IV. DISCUSSION

The continued advancement of medical imaging technology is reaching the stage where procurement of high reso-

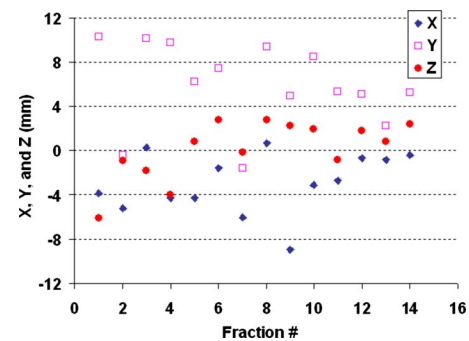


FIG. 7. 3D spatial position of a fiducial in a prostate patient before treatment setup at 14 treatment fractions.

lution anatomical images can be acquired rapidly and with low diagnostic dosages. In addition, utilizing the actual MV treatment beam for imaging has the potential to further reduce diagnostic doses. These various imaging modalities pave the way for real-time IGRT. For the tested projection images, the algorithm demonstrates high successful detection rates together with near real-time speeds. As seen in Figs. 4, 6, and 8, reliance on only R^2 can lead to false positive marker

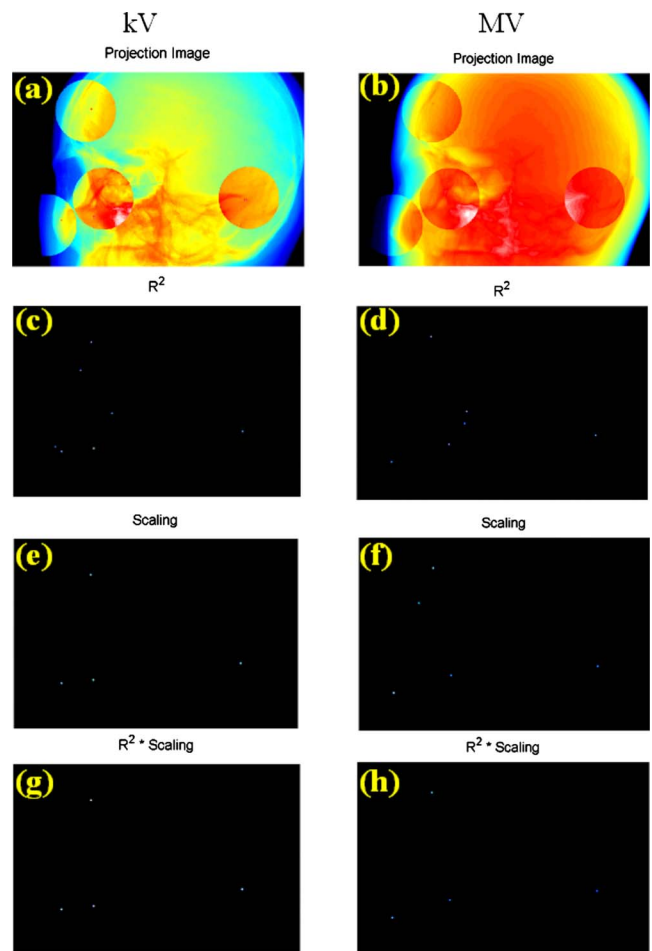


FIG. 8. Analyzing four BBs on kV (left column) and MV (right column) images. (a) and (b) Projection image and the predicted BB ROIs were highlighted, (c) and (d) R^2 results (threshold applied), (e) and (f) scaling factors (threshold applied), and (g) and (h) searching index= $R^2 \cdot$ scaling factor.

detection, especially for MV based projections. The combination of these two criteria resulted in 100% success detection rates for all 3500 images kV or MV projections analyzed. It is important to emphasize that multiple markers could be successfully tracked on MV images.

Fast marker detection is a crucial component of real-time image based tracking. This algorithm quickens the search process in two aspects. First, the orientation and cross-sectional pattern matching for cylindrical markers simplifies the problem due to cylindrical orientation and projection length, which would otherwise require a large number of patterns to be tested.²⁷ Secondly, the search space is significantly reduced by incorporating prior CT marker location knowledge. For a typical 1024×768 image with five cylindrical markers, the processing time is reduced from ~ 1 s to 0.1 s upon incorporating ROI selection. For fewer markers this time would be reduced even further. Generally, spherical marker detection requires approximately half the computational time as compared to detecting the same number of cylindrical markers. In addition, this process is completely automated and does not need any manual location initializations, which is another asset.

Currently, the EPID imager used in this study is hardware limited to a maximum frames-per-sec (fps) rate of ~ 7.5 fps. It is, however, envisioned that as future faster kV and MV detectors become available, the algorithm's detection time can be further reduced. In general, the maximum clinically seen velocity of a tumor is ~ 4.0 cm/s (respiratory based). With the current 7.5 fps imaging speed this would correspond to a between frame marker travel distance of ~ 0.5 cm. This distance is small in comparison to the 2 cm in radius ROI used for detection. To increase computational speeds it is beneficial to reduce the search space defined by the ROI. This could be accomplished by recalculating a new 0.5 cm in radius ROI for each new frame analyzed. The location of this ROI would be centered on the marker's location in the previous frame, thus ensuring that the marker is located within the ROI.

As demonstrated in Fig. 5, the orthogonal location of the kV imager relative the MV imager allows for calculation of the 3D position of the markers from the isocenter. Although this current work consists of retrospective analysis, it is expected that this tracking algorithm can provide real-time 3D tracking in such a system. This is especially convenient on treatment systems pre-equipped with kV and MV onboard imaging equipment as it should only require minimal hardware changes, thus being a cost effective solution for implementing IGRT.

V. CONCLUSION

Fast and reliable localization of implanted metallic fiducials of various shapes has been an important yet challenging problem in kV stereoscopic image guided RT. This problem is further aggravated by much reduced image contrast when MV beam imaging is involved for therapeutic guidance. A new pattern matching algorithm has been proposed to track multiple spherical or cylindrical fiducial markers on both

MV and kV projection images. A completely automated detection, 100% detection efficiency, and fast detection speed (10 frames/sec) enable tracking tumor motion in real-time on a LINAC with both kV and MV imaging systems. This algorithm makes it a suitable candidate for future image-guided radiosurgical procedures.

ACKNOWLEDGMENTS

This work was supported in part by grants from the Department of Defense (Grant No. PC040282), the National Cancer Institute (Grant No. 1R01 CA104205), and the Komen Breast Cancer Foundation (Grant No. BCTR0504071).

^{a)} Author to whom all correspondence should be addressed. Electronic mail: lei@reyes.stanford.edu; Telephone: 650-498-7896; Fax: 650-498-4015.

¹IMRT Collaborative Working Group, "Intensity-modulated radiotherapy: Current status and issues of interest," *Int. J. Radiat. Oncol., Biol., Phys.* **51**, 880–914 (2001).

²J. M. Galvin *et al.*, "Implementing IMRT in clinical practice: A joint document of the American Society for Therapeutic Radiology and Oncology and the American Association of Physicists in Medicine," *Int. J. Radiat. Oncol., Biol., Phys.* **58**, 1616–1634 (2004).

³K. Kitamura *et al.*, "Three-dimensional intrafractional movement of prostate measured during real-time tumor-tracking radiotherapy in supine and prone treatment positions," *Int. J. Radiat. Oncol., Biol., Phys.* **53**, 1117–1123 (2002).

⁴J. M. Crook *et al.*, "Prostate motion during standard radiotherapy as assessed by fiducial markers," *Radiother. Oncol.* **37**, 35–42 (1995).

⁵R. I. Berbeco *et al.*, "Residual motion of lung tumours in end-of-inhale respiratory gated radiotherapy based on external surrogates," *Med. Phys.* **33**, 4149–4156 (2006).

⁶Y. Seppenwoolde *et al.*, "Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **53**, 822–834 (2002).

⁷Y. Xie *et al.*, "Intrafraction motion of the prostate in cyberknife hypofractionated radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* (accepted).

⁸K. L. Zhao *et al.*, "Evaluation of respiratory-induced target motion for esophageal tumors at the gastroesophageal junction," *Radiother. Oncol.* **84**, 283–289 (2007).

⁹H. Shirato *et al.*, "Speed and amplitude of lung tumor motion precisely detected in four-dimensional setup and in real-time tumor-tracking radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **64**, 1229–1236 (2006).

¹⁰L. Xing *et al.*, "Overview of image-guided radiation therapy," *Med. Dosim.* **31**, 91–112 (2006).

¹¹M. J. Murphy, "Tracking moving organs in real time," *Semin. Radiat. Oncol.* **14**, 91–100 (2004).

¹²P. C. Chi *et al.*, "Relation of external surface to internal tumor motion studied with cine CT," *Med. Phys.* **33**, 3116–3123 (2006).

¹³H. Shirato *et al.*, "Real-time tumour-tracking radiotherapy," *Lancet* **353**, 1331–1332 (1999).

¹⁴P. Kupelian *et al.*, "Multi-institutional clinical experience with the Calypso System in localization and continuous, real-time monitoring of the prostate gland during external radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **67**, 1088–1098 (2007).

¹⁵A. L. Boyer *et al.*, "A review of electronic portal imaging devices (EPIDs)," *Med. Phys.* **19**, 1–16 (1992).

¹⁶L. Dong *et al.*, "Verification of radiosurgery target point alignment with an electronic portal imaging device (EPID)," *Med. Phys.* **24**(2), 263–267 (1997).

¹⁷M. G. Herman *et al.*, "Technical aspects of daily online positioning of the prostate for three-dimensional conformal radiotherapy using an electronic portal imaging device," *Int. J. Radiat. Oncol., Biol., Phys.* **57**, 1131–1140 (2003).

¹⁸P. J. Keall *et al.*, "On the use of EPID-based implanted marker tracking for 4D radiotherapy," *Med. Phys.* **31**, 3492–3499 (2004).

¹⁹H. Dehnad *et al.*, "Clinical feasibility study for the use of implanted gold seeds in the prostate as reliable positioning markers during megavoltage

- irradiation," *Radiother. Oncol.* **67**, 295–302 (2003).
- ²⁰A. Nederveen, J. Lagendijk, and P. Hofman, "Detection of fiducial gold markers for automatic on-line megavoltage position verification using a marker extraction kernel (MEK)," *Int. J. Radiat. Oncol., Biol., Phys.* **47**, 1435–1442 (2000).
- ²¹A. J. Nederveen, J. J. Lagendijk, and P. Hofman, "Feasibility of automatic marker detection with an a-Si flat-panel imager," *Phys. Med. Biol.* **46**, 1219–1230 (2001).
- ²²J. Pouliot *et al.*, "(Non)-migration of radiopaque markers used for on-line localization of the prostate with an electronic portal imaging device," *Int. J. Radiat. Oncol., Biol., Phys.* **56**, 862–866 (2003).
- ²³D. C. Schiffner *et al.*, "Daily electronic portal imaging of implanted gold seed fiducials in patients undergoing radiotherapy after radical prostatectomy," *Int. J. Radiat. Oncol., Biol., Phys.* **67**, 610–619 (2007).
- ²⁴L. Smith *et al.*, "Automatic detection of fiducial markers in fluoroscopy images for on-line calibration," *Med. Phys.* **32**, 1521–1523 (2005).
- ²⁵H. Shirato *et al.*, "Physical aspects of a real-time tumor-tracking system for gated radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **48**, 1187–1195 (2000).
- ²⁶T. R. Willoughby *et al.*, "Evaluation of an infrared camera and x-ray system using implanted fiducials in patients with lung tumors for gated radiation therapy," *Int. J. Radiat. Oncol., Biol., Phys.* **66**, 568–575 (2006).
- ²⁷X. Tang, G. C. Sharp, and S. B. Jiang, "Fluoroscopic tracking of multiple implanted fiducial markers using multiple object tracking," *Phys. Med. Biol.* **52**, 4081–4098 (2007).
- ²⁸R. D. Wiersma, W. Mao, and L. Xing, "Combined kV and MV imaging for real-time tracking of implanted fiducial markers," *Med. Phys.* **35**, 1191–1198 (2008).
- ²⁹K. Kitamura *et al.*, "Tumor location, cirrhosis, and surgical history contribute to tumor movement in the liver, as measured during stereotactic irradiation using a real-time tumor-tracking radiotherapy system," *Int. J. Radiat. Oncol., Biol., Phys.* **56**, 221–228 (2003).

PHYSICS CONTRIBUTION

INTRAFACTIONAL MOTION OF THE PROSTATE DURING
HYPOFRACTIONATED RADIOTHERAPYYAOQIN XIE, PH.D.,* DAVID DJAJAPUTRA, PH.D.,* CHRISTOPHER R. KING, PH.D., M.D.,*
SABBIR HOSSAIN, PH.D.,[†] LIJUN MA, PH.D.,[†] AND LEI XING, PH.D.**Department of Radiation Oncology, Stanford University School of Medicine, Stanford, CA; and [†]Department of
Radiation Oncology, University of California San Francisco, San Francisco, CA**Purpose:** To report the characteristics of prostate motion as tracked by the stereoscopic X-ray images of the implanted fiducials during hypofractionated radiotherapy with CyberKnife.**Methods and Materials:** Twenty-one patients with prostate cancer who were treated with CyberKnife between January 2005 and September 2007 were selected for this retrospective study. The CyberKnife uses a stereoscopic X-ray system to obtain the position of the prostate target through the monitoring of implanted gold fiducial markers. If there is a significant deviation, the treatment is paused while the patient is repositioned by moving the couch. The deviations calculated from X-ray images acquired within the time interval between two consecutive couch motions constitute a data set.**Results:** Included in the analysis were 427 data sets and 4,439 time stamps of X-ray images. The mean duration for each data set was 697 sec. At 30 sec, a motion >2 mm exists in about 5% of data sets. The percentage is increased to 8%, 11%, and 14% at 60 sec, 90 sec, and 120 sec, respectively. A similar trend exists for other values of prostate motion.**Conclusions:** With proper monitoring and intervention during treatment, the prostate shifts observed among patients can be kept within the tracking range of the CyberKnife. On average, a sampling rate of ~40 sec between consecutive X-rays is acceptable to ensure submillimeter tracking. However, there is significant movement variation among patients, and a higher sampling rate may be necessary in some patients. © 2008 Elsevier Inc.

CyberKnife, Prostate cancer, Fiducial markers, Real-time tracking.

INTRODUCTION

Recent randomized studies for patients with localized prostate cancer confirm that improved biochemical failure-free survival was achieved by using higher doses of external beam radiotherapy (RT) (1–3). Although a higher dose is good for tumor control, it also carries greater risk of complications to surrounding critical structures, such as the bladder and rectum (4). Because of the inter- and intrafractional motion of the prostate, margin is required when planning a prostate radiotherapy. Knowing the extent of prostate movement during a fractionated, and more important a hypofractionated, treatment is necessary to reduce the treatment margin and facilitate prostate dose escalation (5, 6). A number of techniques have been developed for measuring setup variations and internal organ motion for individual patients from day-to-day and during a treatment fraction (7).

Ultrasound has been a useful tool for prostate target localization (8–10). Fung *et al.* (8) analyzed the data of 7,825 daily fractions of 234 prostate patients and indicated average three-

dimensional (3D) interfractional displacement of about 7.8 mm. Electronic Portal Imaging Device, on-board kV X-ray imaging, or both of implanted fiducials is also widely used for initial setup and interfractional monitoring of the prostate target position (11–17). A recent development in measuring setup variations is the electromagnetic positioning and continuous monitoring system from Calypso Medical Technologies (Seattle, WA) (18–20). The difference between skin marks vs. the Calypso System alignment was found to be >5 mm in vector length in more than 75% of fractions. Displacements >3 mm and 5 mm for cumulative durations of at least 30 sec were observed during 41% and 15% of sessions, respectively.

At our institution, CyberKnife (Accuray, Sunnyvale, CA) has been employed for Phase II hypofractionated treatment of prostate cancer. Through frequent stereoscopic X-ray imaging of implanted fiducials, the CyberKnife provides an effective way to monitor the position of the prostate target during a hypofractionated treatment (21). The system records the center of mass (CM) of implanted fiducials as computed

Reprint requests to: Lei Xing, Ph.D., Stanford University School of Medicine Department of Radiation Oncology, 875 Blake Wilbur Drive, Stanford, CA 94305-5847. Tel: (650) 498-7896; Fax: (650) 498-4015; E-mail: lei@reyes.stanford.edu

Conflict of interest: none.

Acknowledgement—This work was supported in part by grants from the Department of Defense (Grant No. PC040282) and the National Cancer Research Institute (Grant No. 1R01 CA98523 and CA104205).

Received Dec 7, 2007, and in revised form April 10, 2008.
Accepted for publication April 29, 2008.

from each pair of stereoscopic images during each treatment, thus providing a valuable set of data to better understand the intrafractional movement of the prostate. In addition to the technical difference in monitoring the implanted fiducials, a major feature of our data is that the time span of tumor motion monitoring is significantly longer compared with the Calypso data (up to 2,500 sec with a mean duration of about 700 sec vs. 600 sec for Calypso). This study sheds useful insight into features of intrafractional prostate motion and reemphasizes the need for an effective means of compensating the intrafractional prostate movement to ensure adequate dose coverage of the tumor target.

METHODS AND MATERIALS

CyberKnife data acquisition

During hypofractionated prostate radiation treatment, fiducials must be rigidly fixed no more than 5–6 cm relative to a known reference or to the tumor. Any fiducial migration will degrade the accuracy of fiducial-based targeting. Commonly, three fiducials are used for prostate cancer treatment.

The patient setup and treatment delivery process is illustrated in Fig. 1. First, orthogonal X-ray images are acquired before treatment. The system determines the absolute position of the target volume via image-to-digitally reconstructed radiograph registration. The three-dimensional translation and rotation deviation of the target from the planned position is calculated. The deviation is corrected by manu-

ally moving the treatment couch. The treatment starts if the computed shift is less than a preset threshold, 10 mm in general. During treatment, the robot automatically adjusts the incident beam to compensate for the target deviation. The CyberKnife system can perform up to 10 mm translational correction. However, the larger the deviation, the greater is the uncertainty in the accuracy of the robot correction. Therefore it is recommended that the deviation during treatment be kept to a minimum. At Stanford Hospital, a threshold of ~ 5 mm translation is normally used. During the beam delivery, X-ray images are acquired every three nodes, which amounts to about a 40-sec interval. The shift of X-ray images from the planning CT is monitored in real time. If the calculated shift is more than the given threshold, the treatment will be paused, and manual couch movement is required until the shift is below the limit.

Patient selection and prostate motion data analysis

The patients were treated with hypofractionated protocol consisting of five fractions of 7.25 Gy per fraction delivered every other day. In total, 21 prostate cancer patients were treated under the protocol between January 2005 and September 2007 for the study. In our analysis, one fraction can generate more than one data set because the treatment is usually paused a few times to reposition the patient by couch movement. The couch displacement is not kept in treatment log-file, and therefore the data sets before and after the intervening couch movement cannot be joined together without manually writing down the couch shifts during the treatment process.

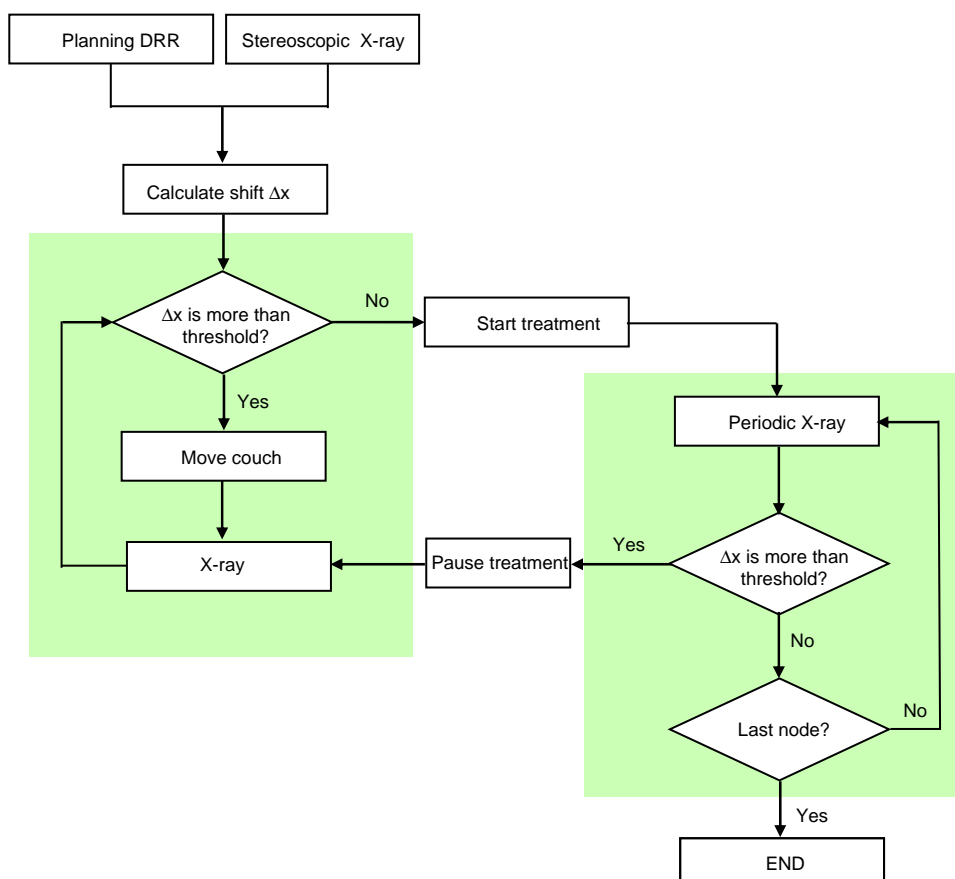


Fig. 1. Flowchart of the patient setup and delivery process in CyberKnife treatment. DRR = digitally reconstructed radiograph.

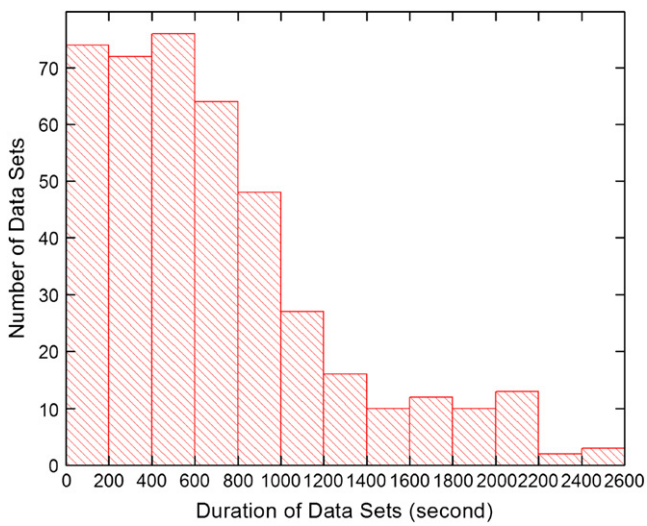


Fig. 2. Histogram of the time span of the studied data sets.

After patient treatment, a log-file containing the CM displacements of the fiducials in anterior–posterior (AP), left–right (LR), and superior–inferior (SI) directions is saved in the CyberKnife control computer and can be readily used for the analysis of prostate movement during the beam delivery process. In addition, files containing the rigid body error (RBE) data of each implanted fiducial are recorded at each timestamp in the Accuray CyberKnife System. The RBE at a timestamp is defined as the distance of a fiducial from its corresponding CT position after the system figures out the best translation and rotation transformation by a rigid registration of the projection images and the CT-generated digitally reconstructed radiographs. For the 21 patients studied here, we first analyzed the statistical characteristics of the collected time duration data sets. This analysis provides valuable information on the average time it takes for the prostate to reach the preset threshold. The CM displacement log-files acquired during the treatment course were then studied and the overall and patient-specific behaviors of prostate displacement were investigated.

Study of prostate deformation

The RBE values mentioned earlier were employed to gain an understanding of prostate deformation. In general, the intrafractional motion of prostate consists of rigid and deformable motions. The rigid motion of the prostate is characterized by the CM displacement of the three implanted fiducials and any potential rotation. In principle, it is possible to estimate the prostate rotation on the basis of the angular change of the fiducial configuration. However, the prostate rotation is generally small and negligible. We have therefore focused our study on translation and deformable motions of the prostate. To a certain extent, the deformation can be described by the RBE values and the correlation between RBE curves of the three implanted fiducials. In an ideal case in which there is no deformation, the RBE should be zero, and the three time-dependent RBE curves would correlate completely. The correlation between three fiducials was computed for four representative cases.

Influence of X-ray image sampling rate

A clinically important question in stereoscopic image-guided prostate treatment is the optimal X-ray imaging frequency. Here the optimal imaging frequency is defined as the frequency that minimizes patient exposure while not missing any significant movement

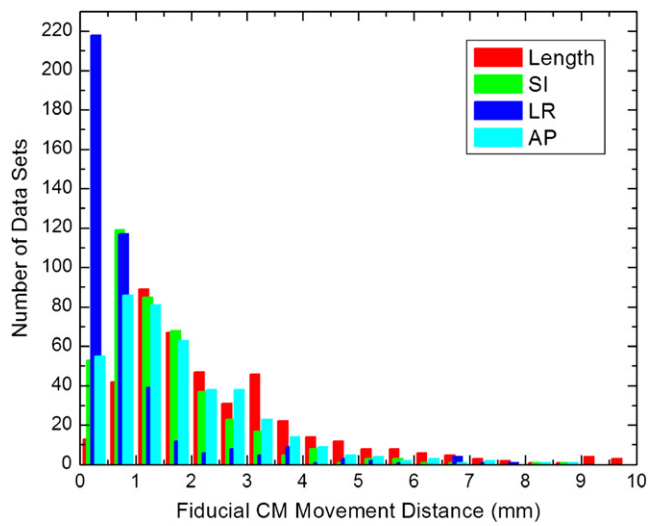


Fig. 3. Histogram of the fiducial center of mass (CM) movement in different directions. AP = anterior–posterior; LR = left–right; SI = superior–inferior.

of the target during the beam-off interval of the imaging X-ray (22). To better understand the issue, we investigated the consequence of down-sampling the X-ray imaging data for one of the patients.

RESULTS

Duration of data sets

For the 21 patients, 4,439 timestamps, which constitute 427 separate data sets, were recorded. Figure 2 shows the histogram of the duration of the 427 data sets. The bin size is 200 sec, and the mean duration is 697 sec. The duration of a data set represents the time for the prostate to move beyond an acceptable level to the therapist (~5 mm shift for prostate at our institution). Thus a shorter duration corresponds to a more prominent prostate movement. The data here suggest that on average it takes 697 sec for the prostate to move beyond 5 mm relative to its planned position.

Overall behavior of intrafractional prostate motion

A useful way to present prostate motion data is to show the histogram of the fiducial CM movement in different directions. As seen in Fig. 3, the prostate is more stable in the LR direction, which is consistent with pelvic and prostate anatomy. Generally, the shift distribution in the SI direction is similar to that in the AP direction. The mean shift in each direction, averaged over all patients, was 1.55 ± 1.28 mm,

Table 1. Statistical characterization of the 427 data sets for each direction (SI, LR and AP) and the vector length of the shift (length)

	SI	LR	AP	Length
Average (mm)	1.55	0.87	1.80	2.61
SD (mm)	1.28	1.17	1.44	1.94

Abbreviations: AP = anterior–posterior; LR = left–right; SI = superior–inferior.

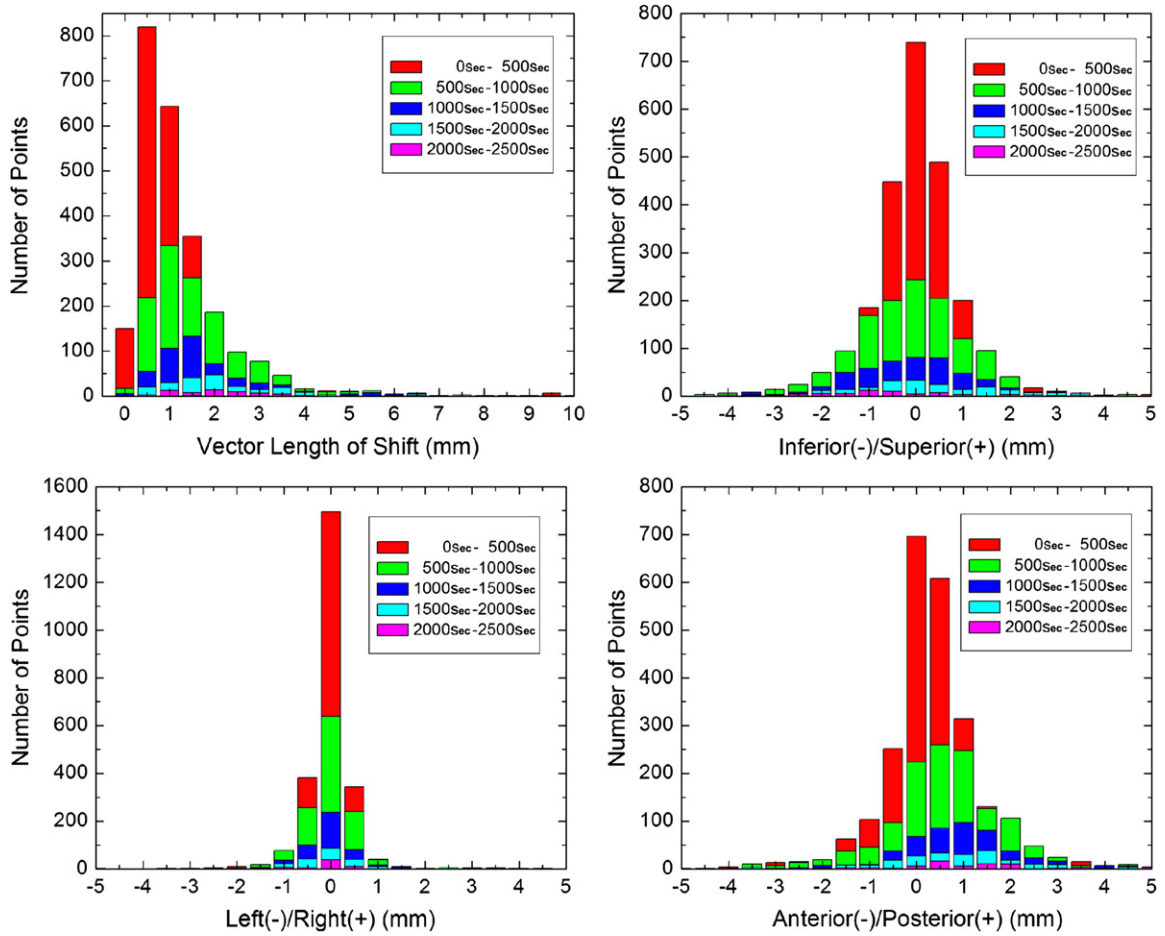


Fig. 4. Histogram of the prostate movement as a function of time duration and shift.

0.87 ± 1.17 mm, 1.80 ± 1.44 mm in SI, LR, and AP directions, respectively. The average vector length of the shift is 2.61 ± 1.94 mm. Table 1 summarizes the statistical characterization of the data for each direction and the vector length of the shift. It should be emphasized that these mean values were specific to the 5-mm threshold used in fiducial tracking because the greater-than-threshold shifts were reset by manual couch adjustments. It therefore does not represent the mean magnitude of prostate motion during a complete treatment fraction.

An alternative way to present the data in Fig. 3 is to illustrate the histogram of the prostate movement as a function of time duration and shift, as shown in Fig. 4. Each color represents a specific time segment. It is clear that as time elapses, the spatial distribution of the prostate becomes increasingly spread out. From these plots, it is also clear that SI and AP movements are similar, whereas the LR curve is more concentrative.

A rolling average (23) of total movement distance was computed in equal time-interval stamps to illustrate further the prostate movement tendency, as shown in Fig. 5. Because the average image acquisition interval is ~ 40 sec and using the Nyquist sampling theory (24), a time interval of 20 sec was used to calculate the rolling average curve. The rolling average window was set to be 120 sec. Thus the average shift

at 0 sec is a result of shifts from 0 to 60 sec, and the average shift at 100 sec represents the contributions from 40 to 160 sec. Because there are not enough data for rolling average calculation at the end of the time duration, the calculation stopped at 1400 sec in Fig. 5. The curve in Fig. 5 represents

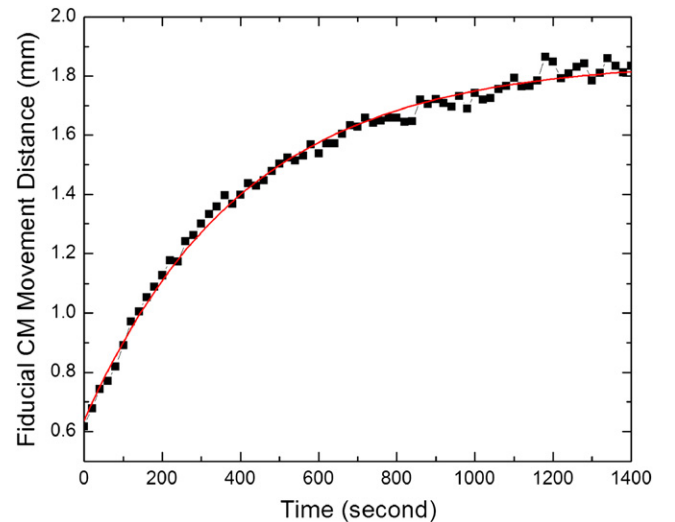


Fig. 5. Rolling averages of prostate center of mass (CM) movement.

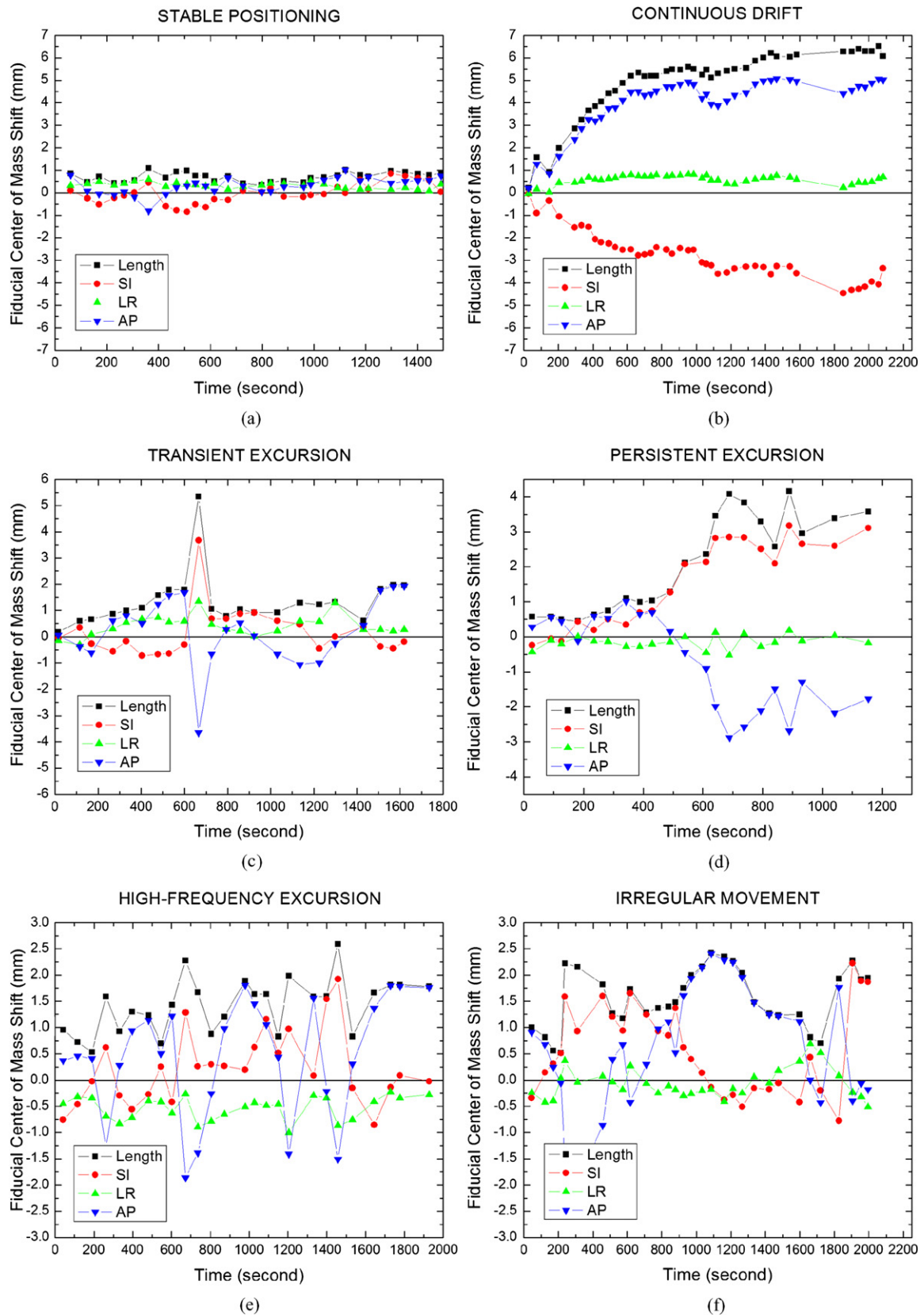


Fig. 6. Patterns of prostate movement: (a) stable target at baseline, (b) continuous drift, (c) transient excursion, (d) persistent excursion, (e) high-frequency excursion, and (f) irregular movement (red: superior/inferior [SI] direction; green: left/right [LR] direction; blue: anterior/posterior [AP] direction; black: vector length of the shift).

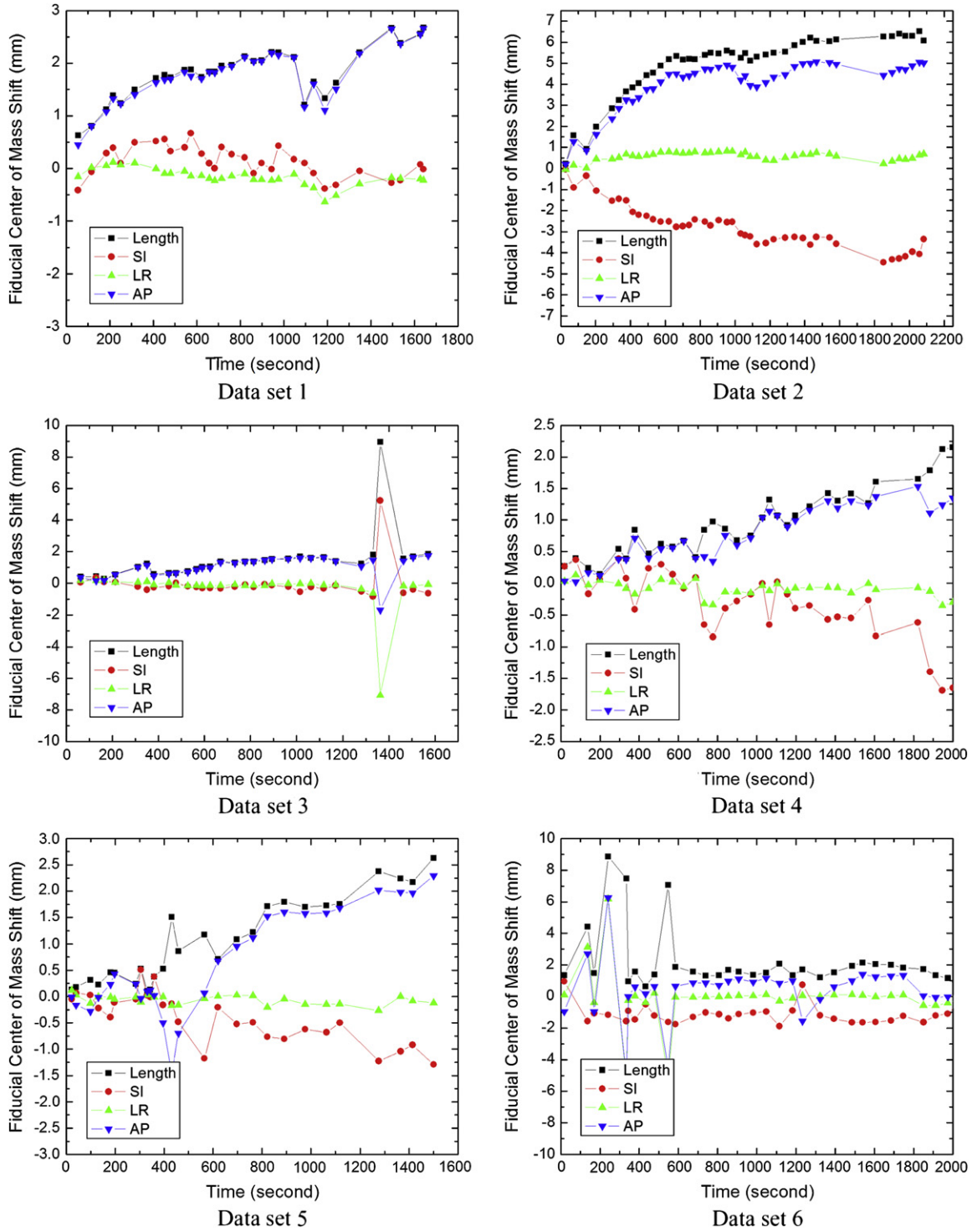


Fig. 7. Prostate movement behaviors for one of the patients in different data sets (red: superior/inferior [SI] direction; green: left/right [LR] direction; blue: anterior/posterior [AP] direction; black: vector length of the shift).

a logarithmic fit of the data. It can be seen that the prostate movement increases against time.

Patient-specific behavior of intrafractional prostate motion

The motion of the CM of the three implanted fiducials is used as a surrogate of prostate motion. This quantity was re-

corded over time and analyzed for the 427 data sets (on average 20 data sets per patient). For illustration, a selection of six typical patterns of motion categories is shown in Fig. 6a–6f. The x axis represents the timestamp, and the y axis is the motion shift. Similar to those observed by Kupelian *et al.* (18), these patterns vary from stable positioning at baseline

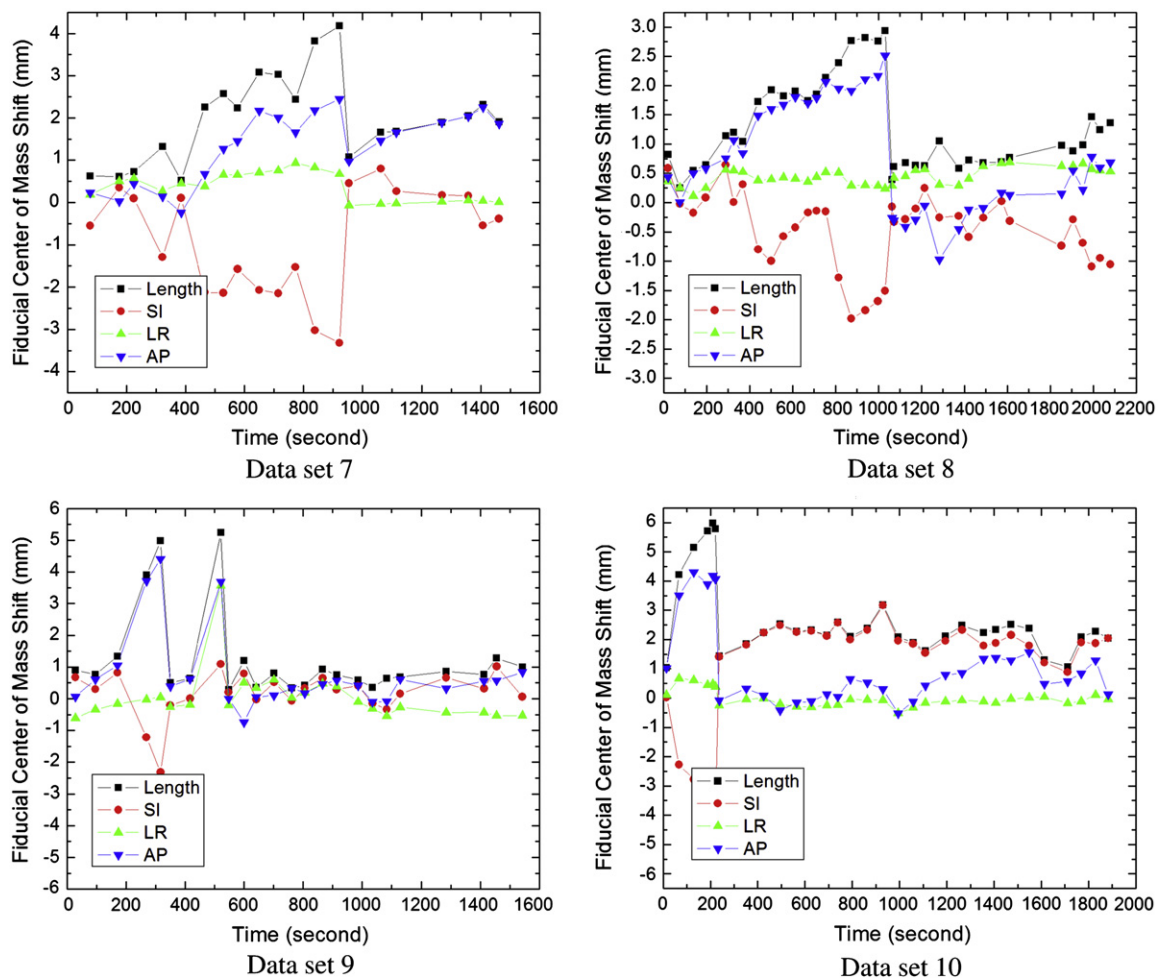


Fig. 7. Continued

(Fig. 6a), continuous drift (Fig. 6b), transient excursion (Fig. 6c), persistent excursion (Fig. 6d), and high-frequency excursion (Fig. 6e). Some patterns are simply too irregular to categorize into any of these classes (Fig. 6f).

It should be noted that for each patient, the prostate movement pattern may change from fraction to fraction or even from data set to data set within the same treatment fraction. Figure 7 shows the prostate movement behavior for one of the patients. This patient received five treatment fractions, and each fraction contains two data sets. Data sets 2, 4, and 5 are the continuous drift; data set 3 is the transient excursion; data sets 1, 7, and 8 consist of two continuous drifts; data sets 6 and 9 are high-frequency excursions followed by a stable positioning at baseline; and data set 10 shows a continuous drift followed by a stable positioning at baseline. These data suggest that the prostate intrafractional motion is somewhat random and does not follow a fixed pattern.

Influence of prostate deformation

Prostate deformation is of a practical concern. In Fig. 8, the RBE curves for each fiducial are plotted for four representative patients. The RBE value is generally within 1.5 mm, in-

dicating that the deformation of the prostate is not a major issue here. Furthermore, as listed in Table 2, the correlation coefficients between RBE curves of the three implanted fiducials were found to be close to 100%. Although the fiducial movement profiles differ among patients, the motion behaviors of three fiducials for each patient are similar, indicating that no prominent deformation occurred in these cases.

Optimal image sampling rate

To understand the influence of the X-ray image sampling rate, we extracted a fraction of movement from one of the patients as shown in the solid line of Fig. 9a. Now suppose that the images are acquired at every other timestamp in the original acquisition schedule; the prostate movement curve would look quite different, as indicated in the dashed line of Fig. 9a. Figure 9b is another example of CM movement curve captured by two sampling rates. The peak in the solid curve revealed by a higher sampling rate disappears when the sampling rate is reduced. It is difficult to guarantee that there is no peak value in a relatively long sample interval. However, more frequent real-time imaging, which provides more accurate correction to the treatment robot, would result

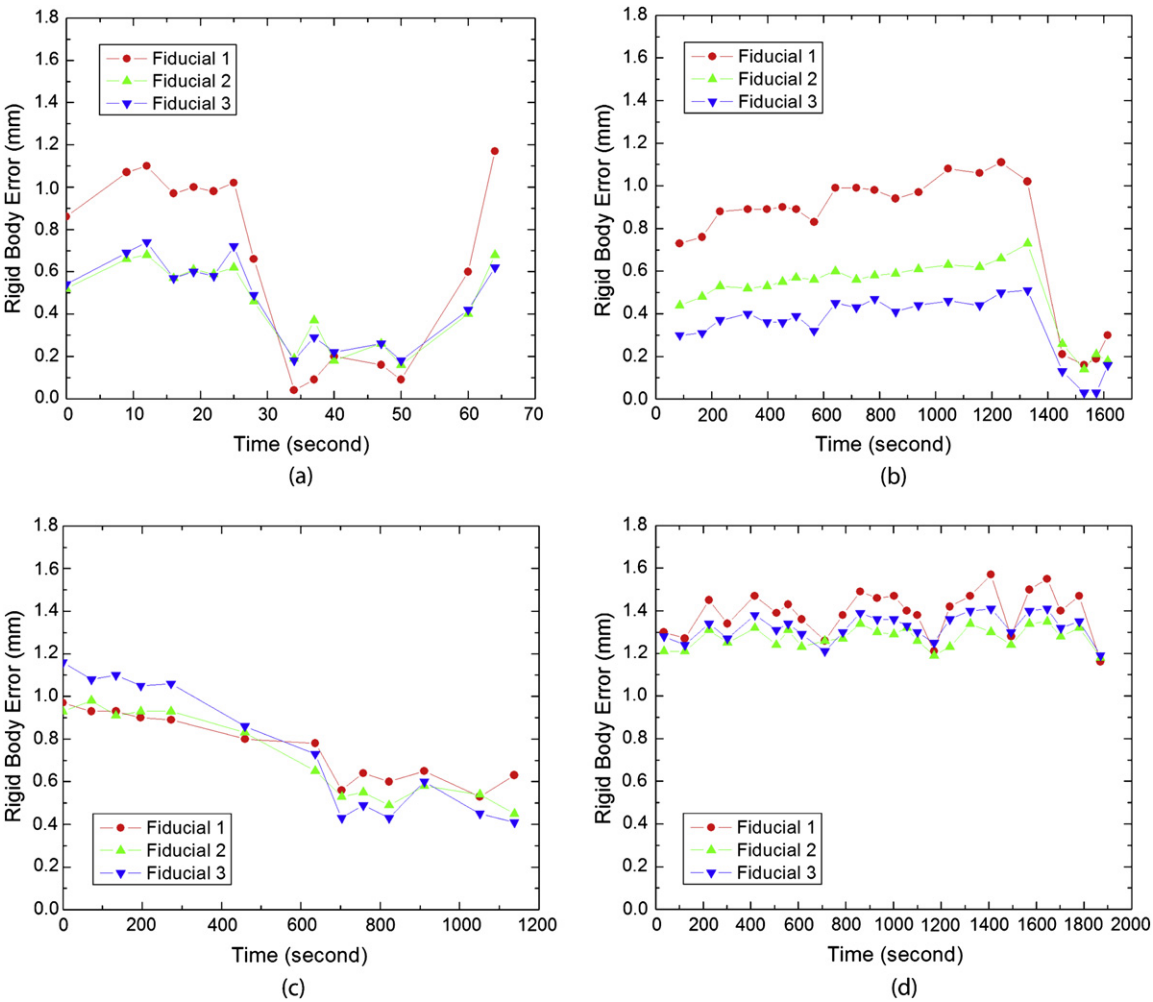


Fig. 8. Rigid body error curve of three fiducials for four representative patients (red: Fiducial 1; green: Fiducial 2; blue: Fiducial 3).

in longer treatment time and an increase in patients’ normal tissue dose.

In reality, a few factors may influence the selection of the sampling rate of the X-ray imaging, including the dose rate, patient-specific characteristics, the fractionation scheme, and so on. Therefore a trade-off between imaging frequency and target position accuracy must be made. A rule of thumb is that the movement of the prostate within the interval of two consecutive images should be less than a prespecified criterion, say 1 or 2 mm. Because of the randomness of prostate movement, this decision can only be made on a statistical basis. The sampling rate should be chosen in such a way that the number of data sets with displacement exceeding a prespeci-

fied motion range should be statistically small. Figure 10 shows a plot of the percentage for the prostate target to move more than 1 mm, 2 mm, 3 mm, 4 mm, 5 mm, ..., at 30 sec, 60 sec, 90 sec, 120 sec.... This figure is useful in helping to find the suitable sampling rate for a prespecified prostate motion range. For example, if motion greater than 2 mm is permissible for less than 5% of the data sets (in other words, if maintaining 95% of the data sets at a motion <2 mm is desirable), from Fig. 10, it is seen that a 30-sec sampling interval should be used. If a 60-sec interval is used, 7.5% of the data sets will have a motion greater than 2 mm. For 90-sec and 120-sec intervals, the percentage with motion greater than 2 mm will increase to 11% and 14%, respectively. For convenience, the percentage of data sets with motion greater than 1 mm to 5 mm is summarized in Table 3 for a few sampling intervals of interest.

Table 2. Correlation coefficients between three fiducials for four representative cases

	Patient 1	Patient 2	Patient 3	Patient 4
Fiducials 1 and 2	96.2%	97.7%	94.9%	97.8%
Fiducials 1 and 3	96.9%	98.2%	98.4%	99.2%
Fiducials 2 and 3	97.6%	97.0%	98.1%	98.8%

DISCUSSION

To cope with the uncertainty in patient setup and tumor target localization, a commonly used method is to add a population-based safety margin to the target and to sensitive

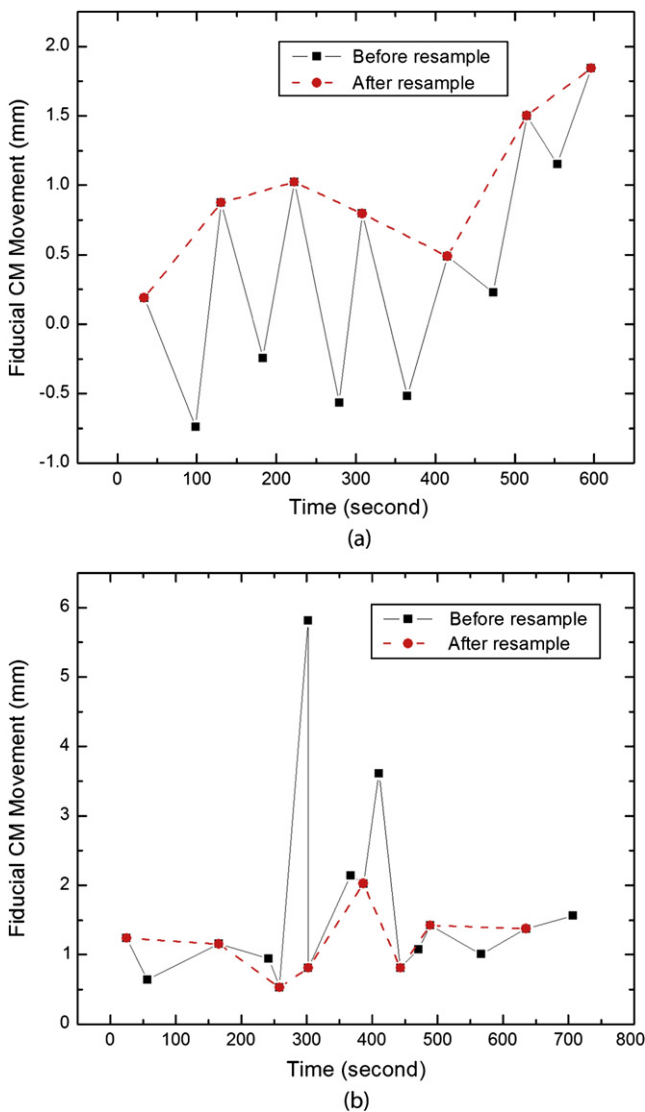


Fig. 9. Prostate movement behaviors depicted by stereoscopic imaging of two sample rates for two patients. (a) Patient 1, (b) Patient 2. CM = center of mass.

adjacent structures to ensure adequate dose coverage, which significantly compromises the success of radiation therapy (25). Because of the proximity of the prostate to the rectum and bladder, a robust strategy in locating the tumor target is necessary if the radiation dose to the prostate is to be escalated to enhance the probability of curing patients without damaging adjacent structures. A detailed knowledge of prostate motion would help us to understand the nature and degree of the adverse influence of the uncertainty and provide guidance in dealing with this issue. The known motion patterns can also be included into inverse planning process to minimize its adverse dosimetric influence.

It is important to bear in mind that the proposed method relies on the assumption that the markers accurately reflect the position of the prostate. In other words, the implanted markers do not migrate significantly within the prostate during the course of treatment. Pouliot *et al.* (26) studied the is-

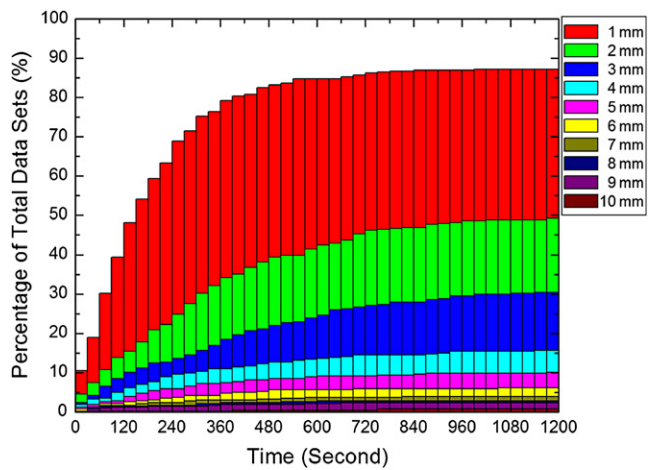


Fig. 10. Percentage of data sets for the prostate target as a function of sample interval and movement threshold.

sue by analyzing the orthogonal portal images of 11 patients. The distances between three markers were determined and monitored over the course of therapy to assess the magnitude of marker migration. The average standard deviation of the distances was found to be 1.3 mm. Similar observations were made by Poggi *et al.* (27), Nichol *et al.* (1), and Shirato *et al.* (28).

The prostate deformation study presented here is estimative in nature, primarily because the RBE and the correlation between RBE fiducial curves for each patient are less quantitative in assessing the organ deformation. However, it is a useful quantity and sheds practical insight onto the problem. A more thorough study based on biomechanical measures of the prostate deformation as a function of time is highly desirable to understand the issue completely.

There are several drawbacks associated with fiducial-based image-guided prostate radiation therapy. Other than the fact that it involves an invasive procedure of fiducial implantation, the fiducial tracking used with CyberKnife or the Calypso system is limited to “rigid” tumors. Although our data indicated that the deformation of the prostate gland during the hypofractionated treatment is small and the CM of the three implanted fiducials can be used to describe the prostate position, it is important to remember that tracking the prostate is only part of the overall task in prostate radiation therapy. In reality, tracking the motion of various adjacent sensitive structures represents the other side of the coin and is also

Table 3. Percentage of data sets having a movement threshold from 1 to 5 mm for a few sampling intervals of interest

Sampling interval	Prostate Movement Threshold				
	1 mm	2 mm	3 mm	4 mm	5 mm
30	10.5%	4.4%	2.3%	1.9%	1.2%
60	19.0%	7.5%	4.2%	3.2%	1.9%
90	30.2%	10.8%	6.6%	3.5%	2.6%
120	39.3%	13.8%	8.4%	4.9%	2.8%

of critical importance to the success of image-guided radiotherapy. Therefore, knowing the spatial location and geometric shapes of the sensitive structures is critical to customizing the dose distribution to maximize the dose to the target while sparing the adjacent sensitive structures. On a fundamental level, the motion of the prostate target is often caused by the motion or physiological change of sensitive structures.

This study, and most if not all similar studies, have focused on prostate-only treatment. Clinically, the treatment of intermediate- and high-risk prostate cancer often involves the irradiation of seminal vesicles and regional lymph nodes. The implanted fiducials in these cases are less helpful in locating the seminal vesicles and pelvic nodes. A better imaging method capable of providing three-dimensional anatomy would be highly desirable. On-board cone-beam CT (CBCT) has recently become available to provide volumetric information of a patient in the treatment position (29). It holds promises for improved target localization and irradiation dose verification (30). CBCT is valuable in providing 3D or even 4D patient model before treatment and affords a useful solution to reduce the adverse effect of interfractional organ motion (22, 31–34). However, acquiring real-time patient geometric information during radiation delivery using an on-board imaging device is still impractical. A combined use of pretreatment patient geometric model derived from 3D/4D

CBCT and real-time stereoscopic X-ray projection data may be useful to estimate the location of target organs and adjacent sensitive structures. This investigation is still in progress and will be reported in the future.

CONCLUSION

Intrafractional organ motion has long been recognized as one of the major limiting factors of prostate dose escalation in conformal radiation therapy. A detailed knowledge of prostate motion would help us to understand the nature and degree of the adverse influence of such motion and provide guidance in dealing with it. Known motion patterns can also be included in the inverse planning process to minimize the adverse dosimetric influence of motion. Our study shows the importance of real-time image guidance and motion-compensation techniques such as the robotic linear accelerator used in CyberKnife during hypofractionated prostate radiation treatment. Given the magnitude and random nature of prostate motion as well as recent technical advancements in various related fields, real-time monitoring of prostate position to compensate for the motion should be part of future prostate radiation therapy to ensure adequate dose coverage of the target while maintaining adequate sparing of adjacent structures.

REFERENCES

- Nichol AM, Brock KK, Lockwood GA, *et al.* A magnetic resonance imaging study of prostate deformation relative to implanted gold fiducial markers. *Int J Radiat Oncol Biol Phys* 2007;67:48–56.
- Zietman AL, DeSilvio ML, Slater JD, *et al.* Comparison of conventional-dose vs high-dose conformal radiation therapy in clinically localized adenocarcinoma of the prostate: A randomized controlled trial. *JAMA* 2005;294:1233–1239.
- Pollack A, Zagars GK, Starkschall G, *et al.* Prostate cancer radiation dose response: Results of the M.D. Anderson Phase III randomized trial. *Int J Radiat Oncol Biol Phys* 2002;53:1097–1105.
- King CR, Lehmann J, Adler JR, *et al.* CyberKnife radiotherapy for localized prostate cancer: Rationale and technical feasibility. *Technol Cancer Res Treat* 2003;2:25–30.
- Van den Heuvel F, Powell T, Seppi E, *et al.* Independent verification of ultrasound based image-guided radiation treatment, using electronic portal imaging and implanted gold markers. *Med Phys* 2003;30:2878–2887.
- Chen J, Lee RJ, Handrahan D, *et al.* Intensity-modulated radiotherapy using implanted fiducial markers with daily portal imaging: Assessment of prostate organ motion. *Int J Radiat Oncol Biol Phys* 2007;68:912–919.
- Langen KM, Jones DT. Organ motion and its management. *Int J Radiat Oncol Biol Phys* 2001;50:265–278.
- Fung AY, Ayyangar KM, Djajaputra D, *et al.* Ultrasound-based guidance of intensity-modulated radiation therapy. *Med Dosim* 2006;31:20–29.
- Chandra A, Dong L, Huang E, *et al.* Experience of ultrasound-based daily prostate localization. *Int J Radiat Oncol Biol Phys* 2003;56:436–447.
- Langen KM, Pouliot J, Anezinos C, *et al.* Evaluation of ultrasound-based prostate localization for image-guided radiotherapy. *Int J Radiat Oncol Biol Phys* 2003;57:635–644.
- Langen KM, Meeks SL, Poole DO, *et al.* The use of megavoltage CT (MVCT) images for dose recomputations. *Phys Med Biol* 2005;50:4259–4276.
- Meeks SL, Harmon JF Jr., Langen KM, *et al.* Performance characterization of megavoltage computed tomography imaging on a helical tomotherapy unit. *Med Phys* 2005;32:2673–2681.
- Pouliot J, Bani-Hashemi A, Chen J, *et al.* Low-dose megavoltage cone-beam CT for radiation therapy. *Int J Radiat Oncol Biol Phys* 2005;61:552–560.
- Litzenberg DW, Balter JM, Lam KL, *et al.* Retrospective analysis of prostate cancer patients with implanted gold markers using off-line and adaptive therapy protocols. *Int J Radiat Oncol Biol Phys* 2005;63:123–133.
- Balter JM, Lam KL, McGinn CJ, *et al.* Improvement of CT-based treatment-planning models of abdominal targets using static exhale imaging. *Int J Radiat Oncol Biol Phys* 1998;41:939–943.
- Wong JR, Grimm L, Uematsu M, *et al.* Image-guided radiotherapy for prostate cancer by CT-linear accelerator combination: prostate movements and dosimetric considerations. *Int J Radiat Oncol Biol Phys* 2005;61:561–569.
- Poulsen PR, Muren LP, Hoyer M. Residual set-up errors and margins in on-line image-guided prostate localization in radiotherapy. *Radiother Oncol* 2007;85:201–206.
- Kupelian P, Willoughby T, Mahadevan A, *et al.* Multi-institutional clinical experience with the Calypso System in localization and continuous, real-time monitoring of the prostate gland during external radiotherapy. *Int J Radiat Oncol Biol Phys* 2007;67:1088–1098.
- Balter JM, Wright JN, Newell LJ, *et al.* Accuracy of a wireless localization system for radiotherapy. *Int J Radiat Oncol Biol Phys* 2005;61:933–937.

20. Willoughby TR, Kupelian PA, Pouliot J, *et al.* Target localization and real-time tracking using the Calypso 4D localization system in patients with localized prostate cancer. *Int J Radiat Oncol Biol Phys* 2006;65:528–534.
21. Dieterich S. Dynamic tracking of moving tumors in stereotactic radiosurgery. In: Mould RF, editor. *Robotic radiosurgery*, Vol 1. Sunnyvale, CA: The CyberKnife Society; 2005. p. 51–63.
22. Xing L, Thorndyke B, Schreiber E, *et al.* Overview of image-guided radiation therapy. *Med Dosim* 2006;31:91–112.
23. Riley JD. Evaluation of travel time estimates derived from automatic vehicle identification tags. Blacksburg, VA: Department of Civil Engineering, Virginia Polytechnic Institute; 1999. p. 98–102.
24. Jones NB, Watson M. Digital signal processing: principles, devices, and applications. London: Institution of Engineering and Technology; 1990.
25. Rasch C, Steenbakkers R, van Herk M. Target definition in prostate, head, and neck. *Semin Radiat Oncol* 2005;15:136–145.
26. Pouliot J, Aubin M, Langen KM, *et al.* (Non)-migration of radiopaque markers used for on-line localization of the prostate with an electronic portal imaging device. *Int J Radiat Oncol Biol Phys* 2003;56:862–866.
27. Poggi MM, Gant DA, Sewchand W, *et al.* Marker seed migration in prostate localization. *Int J Radiat Oncol Biol Phys* 2003;56:1248–1251.
28. Shirato H, Harada T, Harabayashi T, *et al.* Feasibility of insertion/implantation of 2.0-mm-diameter gold internal fiducial markers for precise setup and real-time tumor tracking in radiotherapy. *Int J Radiat Oncol Biol Phys* 2003;56:240–247.
29. Jaffray DA, Siewerdsen JH, Wong JW, *et al.* Flat-panel cone-beam computed tomography for image-guided radiation therapy. *Int J Radiat Oncol Biol Phys* 2002;53:1337–1349.
30. Yang Y, Schreiber E, Li T, *et al.* Evaluation of on-board kV cone beam CT (CBCT)-based dose calculation. *Phys Med Biol* 2007;52:685–705.
31. Li T, Koong A, Xing L. Enhanced 4D cone-beam CT with interphase motion model. *Med Phys* 2007;34:3688–3695.
32. Li T, Schreiber E, Yang Y, *et al.* Motion correction for improved target localization with on-board cone-beam computed tomography. *Phys Med Biol* 2006;51:253–267.
33. Li T, Xing L. Optimizing 4D cone-beam CT acquisition protocol for external beam radiotherapy. *Int J Radiat Oncol Biol Phys* 2007;67:1211–1219.
34. Sonke JJ, Zijp L, Remeijer P, *et al.* Respiratory correlated cone beam CT. *Med Phys* 2005;32:1176–1186.

Feature-based rectal contour propagation from planning CT to cone beam CT

Yaoqin Xie, Ming Chao, Percy Lee, and Lei Xing^{a)}

Department of Radiation Oncology, Stanford University School of Medicine, Stanford, California 94305-5847

(Received 13 December 2007; revised 13 May 2008; accepted for publication 4 August 2008; published 12 September 2008)

The purpose of this work is to develop a novel feature-based registration strategy to automatically map the rectal contours from planning computed tomography (CT) (pCT) to cone beam CT (CBCT). The rectal contours were manually outlined on the pCT. A narrow band with the outlined contour as its interior surface was then constructed, so that we can exclude the volume inside the rectum in the registration process. The corresponding contour in the CBCT was found by using a feature-based registration algorithm, which consists of two steps: (1) automatically searching for control points in the pCT and CBCT based on the features of the surrounding tissue and matching the homologous control points using the scale invariance feature transformation; and (2) using the control points for a thin plate spline transformation to warp the narrow band and mapping the corresponding contours from pCT to CBCT. The proposed contour propagation technique is applied to digital phantoms and clinical cases and, in all cases, the contour mapping results are found to be clinically acceptable. For clinical cases, the method yielded satisfactory results even when there were significant rectal content changes between the pCT and CBCT scans. As a consequence, the accordance between the rectal volumes after deformable registration and the manually segmented rectum was found to be more than 90%. The proposed technique provides a powerful tool for adaptive radiotherapy of prostate, rectal, and gynecological cancers in the future. © 2008 American Association of Physicists in Medicine. [DOI: [10.1118/1.2975230](https://doi.org/10.1118/1.2975230)]

Key words: image guided radiation therapy (IGRT), image registration, deformable model, segmentation, scale invariance feature transformation (SIFT)

I. INTRODUCTION

Patients treated with radiotherapy for cancers such as prostate, rectal, and gynecological cancers experience large day-to-day changes in their rectal volumes due to motion, distension, and filling. Due to variations in the image content, an exact correspondence between two image sets acquired at different time points may not exist. Thus, any deformable model relying on the use of information contained in the entire image may not be adequate in dealing with these patients. The artifacts-induced disjoint between the images also makes the autopropagation of contours outlined in one set of images to another highly difficult with conventional strategies. With continued enthusiasm for adaptive radiotherapy, the ability to reliably and efficiently map the rectum outlined in the planning computed tomography (CT) (pCT) to the on-treatment cone beam CT (CBCT) images now becomes a bottleneck and needs to be resolved in order for many patients with cancer within the pelvis to benefit from the novel adaptive replanning strategy.^{1,2}

The issue of rectal motion and deformation in conformal radiation therapy is described in various publications. Lee *et al.* evaluated the CBCT as a tool to quantify the accuracy and precision of a simulated IMRT treatment delivery model for rectal cancer when rectal motion due to filling and deformation was taken into account.³ The mean deformation variation of 0.71 and 0.94 cm in the LAT and AP directions was

reported. Foskey *et al.* shrank the rectal gas region to a virtual point in order to make the correspondence of the rectal volumes in two sets of images.⁴ Gao *et al.* used an automatic image intensity modification procedure to create artificial gas pockets in the pCT images.⁵ The major drawbacks of these types of approaches are the artificial introduction of image features within the rectal volume and the potentially inaccurate association of the artificial image features. As a consequence, the accordance between the rectal volumes after deformable registration and the manually segmented rectum was found to be less than 80%.

In this work, we propose to use the image information in the neighborhood outside the rectal wall as the driving force to guide the rectal contour propagation from the pCT to CBCT. Because the content in the region outside the rectal wall should be conserved, regardless of any changes in the rectal filling and distension, this strategy seems to be physically sensible. Coupled with a powerful feature-based deformable registration model, which identifies homologous tissue features shared by the pCT and CBCT images, the novel approach captures the key issues of the system and provides a natural solution to the above stated problem. Application of the proposed algorithm to a number of digital phantoms and clinical cases demonstrates that the technique is accurate and robust and may be useful for future adaptive therapy planning.

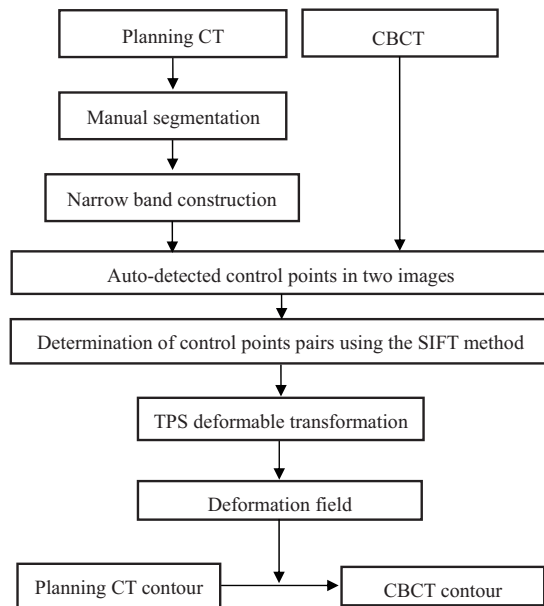


FIG. 1. Overall process of rectal contour propagation.

II. METHODS AND MATERIALS

II.A. Software platform

The proposed contour mapping algorithm was implemented using the Insight Toolkit⁶ and the Visualization Toolkit (VTK),⁷ which are open source cross-platform C++ software toolkits sponsored by the National Library of Medicine. They are freely available for research purposes (see Refs. 34 and 35). ITK provides various basic algorithms to perform registration and segmentation for medical images. The programs contained in ITK are highly extendable, making it an ideal platform for development of image registration and processing techniques. VTK is primarily used for image visualization (including contours).

II.B. Narrow band construction

Inconsistency in rectal contents between two input image sets could severely reduce the performance of a deformable registration algorithm. Coregistering an empty rectum without bowel gas to a rectum filled with bowel gas using any deformable model could be problematic, for example. A natural strategy is to exclude the volume inside the rectal wall. In practice, the template rectal contour in the pCT image has been manually contoured as a part of the routine treatment planning process, thus making it a straightforward matter to exclude the volume inside the rectal wall. Figure 1 shows the proposed contour mapping process. After manual segmentation on the pCT, a narrow band as sketched in Fig. 2 is constructed with the manually segmented rectum representing the inner surface of the band. On an axial slice, the contour has a polygon shape and the vertices of the polygon form the basis for constructing the narrow band. The distance between the neighboring vertices on the contour is typically 2–10 mm depending on the shape of the contour. In generating the narrow band, we first create squares with side

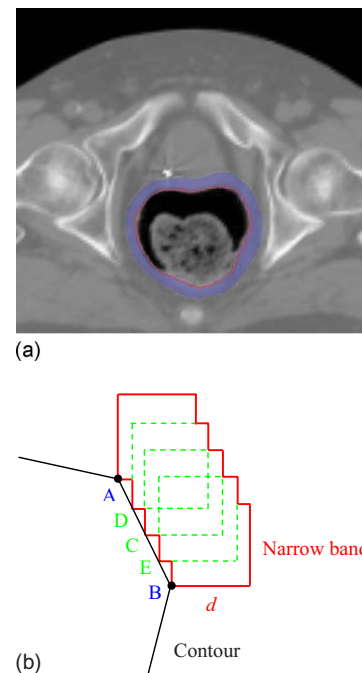


FIG. 2. A sketch of narrow band. (a) A narrow band image surrounding a manually segmented rectal contour and (b) a narrow band construction is illustrated for two vertex points A and B.

length of d for each vertex, as depicted by points A and B in Fig. 2(b). In order to obtain a smooth band, between A and B three more squares, cornered at points C, D, and E, are inserted. Point C is chosen to be the middle point between A and B. Point D is the point between A and C, and point E is the point between B and C. More interpolated vertex points can be similarly introduced to obtain a smooth band. The principle of the narrow band diameter selection is to exclude most bony structures outside the narrow band, since the bony structures are rigid and heavily affect the control point selection. Meanwhile, the generated narrow band can capture sufficient information to drive the finding of its counterpart in the subsequent CBCT. In general, the size of the squares is therefore within 1 cm, so that the diameter of the narrow band is within 1.5 cm.

The narrow band in our approach is used as a compact representation of the rectal surface. As will be detailed in the next subsection, a feature-based deformable registration algorithm is employed to find the correspondence of the band in the CBCT images. Upon successful registration, the deformation field is utilized to propagate the pCT contour to the CBCT. Because only the image features outside the rectum are used, a narrow band shown in Fig. 2 permits us to take advantage of the regional information inside the narrow band yet avoiding the nuisance of rectum/bladder filling.

II.C. Feature-based warping of the narrow band

As illustrated in Fig. 1, the process of contour mapping is to warp the narrow band constructed above in such a way that its best match in the CBCT images is found. Mathematically, this constitutes an optimization problem, in which a

group of transformation parameters transform the points within the band in the pCT to their corresponding points in the CBCT. The input to the contour mapping software includes the narrow band and the CBCT images, which are described by the image intensity distributions $I_a(\mathbf{x})$ and $I_b(\mathbf{x})$, respectively.

To find the transformation matrix, $T(\mathbf{x})$, that maps an arbitrary point in the band to the corresponding point in the CBCT images (or *vice versa*), a thin plate spline (TPS) deformable model is employed. But other models should also be applicable to model the deformation of the band. We automate the control point selection by using the scale invariance feature transformation (SIFT) tissue feature searching (see next subsection for details). Roughly, 300 control points are selected based on the prominent tissue features.

The detailed description of the TPS transformation can be found in Ref. 8. For two-dimensional (2D) images, a weighting vector $W=(w_1, w_2, \dots, w_n)$ and the coefficients a_1, a_u, a_v are computed from a series of matrices which are constructed using n pairs of selected control points in the fixed image (x_i, y_i) and in the moving image (u_i, v_i) , respectively. The function transforming a pixel coordinate in the moving image to a new coordinate in the fixed image is defined as

$$f(u', v') = a_1 + a_u u + a_v v + \sum_{i=0}^n w_i U(|p_i - (u, v)|), \quad (1)$$

where p_i is the control points coordinate in the fixed image and U is a basis function to measure the distance.

II.D. SIFT

The feature-based deformable registration is an essential part of the proposed contour mapping process. Here, we automate the control point selection by using the SIFT-based tissue feature searching. Because of the efficient use of *a priori* system knowledge, the approach greatly enhances the robustness of the narrow band warping algorithm.

The SIFT method was introduced by Lowe to characterize the local tissue features. The method utilizes both image intensity and local gradient information to characterize the neighborhood property of a point.⁹ The algorithm includes scale-space extrema detection, control point localization, orientation assignment, and control point descriptor. In 2D cases, for example, the method uses the orientation histograms of the four quadrants surrounding a point (containing 64 pixels) to characterize the inherent tissue feature of the point (see Fig. 3). To obtain the histogram for a quadrant, as illustrated in Fig. 3, the gradient of each of the 16 pixels in a quadrant is computed. An eight-bin histogram, with first bin representing the number of pixels whose gradients fall between 0° and 45° , and so forth, is then constructed. For illustration, the histogram of each of the four quadrants is displayed schematically in the right panel of Fig. 3 as an eight-vector plot. In total, 32 vectors are calculated in 2D case. In extending the SIFT method from 2D to three dimensional (3D), total of 192 vectors are needed. These vectors

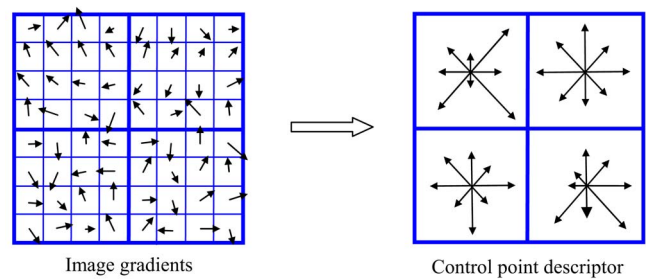


FIG. 3. A sketch of orientation histogram in SIFT method. The gradient of each of the 16 pixels in a quadrant is computed. An eight-bin histogram, with first bin representing the number of pixels whose gradients fall between 0° and 45° , and so forth, is then constructed. The histogram of each of the four quadrants is displayed schematically in the right panel as an eight-vector plot.

represent the local feature and serve as a signature of the point. The SIFT descriptor is considered as one of the most effective descriptors currently available.^{10,11}

Theoretically, the SIFT descriptor can be computed for each voxel in an image. However, this is computationally expensive. The commonly used sampling strategy is to compute the descriptor every 2–3 voxels in x , y , and z directions. After the SIFT descriptors are computed in both input images, the points having the most similar SIFT descriptors in the two images are then identified. For a given point, indexed by n , in the pCT image, the least-squares difference of the SIFT descriptor of the point and that of a potential association point n' in the CBCT, $S_{n,n'}$, is first computed according to

$$S_{n,n'} = \sqrt{\sum_{\alpha=1}^k |(\nabla I_n)_\alpha - (\nabla I_{n'})_\alpha|^2}, \quad (2)$$

where I represents the image intensity. α indexes the bins of the SIFT histogram of a point and the summation over α runs from 1 to 32 for the 2D case, and 1 to 192 for the 3D case. Typically, about 1000 SIFT descriptors n, n' are computed in the narrow band in the pCT and CBCT, respectively. It is unnecessary to determine $S_{n,n'}$ for all possible combinations n, n' , which may dramatically increase the calculation time. We use a specific search radius to control the number of $S_{n,n'}$ calculation. The mapping results are more accurate with larger search radius, however, the calculation time of SIFT mapping becomes longer. After $S_{n,n'}$ is computed, two points n'_1 and n'_2 that have the least histogram difference with point n are identified. If the ratio (for convenience, the ratio is referred to as the κ ratio hereafter) of these two values is less than 80%, the point that has the least S value is chosen tentatively as the correspondence of the point n , otherwise, no association is made for the point. The κ ratio varies between 0 and 1 and is an empirical measure of feature correspondence between two images. The lower the κ ratio, the “stronger” the association of the two feature points on pCT and CBCT. Because of the inherent difference in the textures of the involved organs, the determination of the κ ratio may be organ specific. Typically, it is determined by a tradeoff be-

tween the number of associated point pairs and the reliability of the associations. For a bone, the feature is clear and hundreds point pairs can be associated under a threshold of 50%. On the other hand, for the rectum, the feature is not as obvious as bone. If we still use this low κ ratio, the number of association pairs may be very limited. In this situation, a higher threshold, say 80%, is usually used to increase the number of associated point pairs.

To further increase the accuracy of feature point association, a bidirectional mapping strategy is developed based on the fact that if a point in the pCT is mapped correctly to the CBCT, it will be default to be mapped back to the original point in the pCT when an inverse map is applied to the corresponding point in the CBCT. Therefore, after the original association of feature points as described above, the mapped points in CBCT are inversely coregistered to the pCT. If the correspondence still exists, the associated point pair is labeled a match. Otherwise, they are considered as a mismatch and deleted from the list of correspondence points. Upon the association of the feature points, the associated points are employed as control points. The control point in pCT and CBCT corresponds each other, thus the numbers of control points in the two input images are the same. It was noticed that, when the CBCT region of interest (ROI) is expanded, the increase of feature point generation does not affect the control point association and final contour mapping. The coordinates of an arbitrary point on the contour in CBCT are obtained by interpolating the displacement vectors of the control points using TPS transformation after the control point association is established.

II.E. Evaluation of the models using digital phantom and existing patient data

The performance of the above model is evaluated by a number of 2D digital phantoms and archived clinical cases. In the digital phantom experiments, two deformations are introduced. A virtue of this approach is that the “ground truth” solutions exist and the transformation matrices are known, thus making the evaluation straightforward. The mathematical transformations used to deform the phantom are generated using a formula¹²

$$x'(x,y) = (1 + b \cos m\theta)x, \quad (3)$$

$$y'(x,y) = (1 + b \cos m\theta)y. \quad (4)$$

Here, $\theta = \tan^{-1} y/x$. Two parameters, m and b , are used to characterize a deformation. Generally, they describe the complexity and magnitude of a deformation, respectively. The contour outlined in the original image is then mapped to the deformed image. The accuracy of the contour mapping calculation is assessed by comparing directly with the deformable mapping from the known transformation matrix.

Contour propagation from pCT to CBCT is studied by using three prostate cancer patients and two rectal cancer cases. The pCT is acquired with a GE Discovery-ST CT scanner (GE Medical System, Milwaukee, WI) approximately two weeks prior to the initiation of the radiotherapy.

The on-treatment CBCT images are acquired using the Varian Trilogy™ (Varian Medical Systems, Palo Alto, CA). Each slice of pCT or CBCT is discretized into 512×512 voxels. The images are transferred through DICOM to a high-performance personal computer with a Xeon (3.6 GHz) processor for image processing. The manually outlined contours in the pCT images are mapped to CBCT images using the proposed technique. For the cases studied here, the accordance between the rectal volumes after deformable registration and the manually segmented rectum is employed to assess the success of the proposed algorithm.

To quantitatively evaluate the result of contour propagation, the accordance value between the automapped contour and manually outlined contours were calculated. In general, suppose A and B are two contours, the accordance value r is defined as

$$r = \frac{V_A \cap V_B}{V_A \cup V_B}, \quad (5)$$

here, V is the containing volume of A or B .

III. RESULTS

III.A. 2D digital phantom experiment

The proposed algorithm is first tested using a 2D digital phantom [Fig. 4(a)] with two intentionally introduced deformations of the image shown in Figs. 4(b) and 4(c), respectively. The rectal contour is manually outlined and shown in Fig. 4(a). The deformation shown in Figs. 4(b) and 4(c) are obtained by setting the parameters b and m in Eqs. (3) and (4) to $(b=2, m=2)$ and $(b=2, m=3)$, respectively. The curves close to the interior surface of the rectum in Figs. 4(b) and 4(c) represent the automapped contour. For comparison, the original contour in Fig. 4(a) is also mapped rigidly to Figs. 4(b) and 4(c). Overall, the mapped contours can capture the main features of the two dramatic deformations, and conform to the boundary of the rectum in both cases.

In obtaining the result shown in Fig. 4(b), a total of 200 control points were identified by the bidirectional SIFT calculation as described in method. Note that the bony structure in the image has been excluded in this calculation by setting an intensity threshold of 300 CT number. In this way, any unphysical bony structure deformation is avoided. For clarity, a selection of the SIFT-identified control point associations are displayed in Fig. 5. The superior contour represents the superior surface of narrow band. The total number of control points identified here are far more than that commonly used in TPS calculation,¹³ allowing an improved deformable warping of the narrow band. We should notice that the control points 2, 3, 4, 5, and 9 in Fig. 5 are relatively far away from the rectum wall compared to control points 6 and 10. Since the TPS interpolation is used after SIFT mapping, every control point including points 2, 3, 4, 5, and 9 will affect the deformable warping and therefore the contour shape, although the weights of points 2, 3, 4, 5, and 9 are smaller than points 6 and 10. The displacement field derived by using TPS method is shown in Fig. 6(a). For comparison,

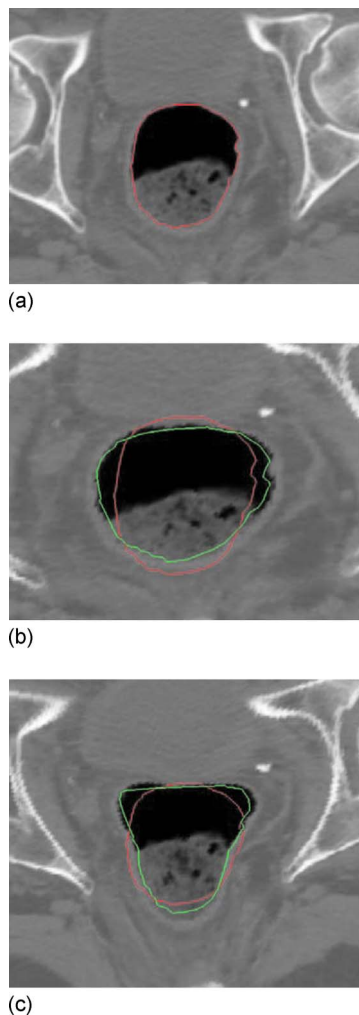


FIG. 4. Rectal contour propagation from the 2D pCT slice to two dramatically deformed images. (a) Original contour, (b) and; (c) its optimal mapping in the two deformed images.

the known displacement field from Eqs. (3) and (4) is plotted in Fig. 6(b). The subtraction between the TPS-derived displacement field and the known field is shown in Fig. 6(c). It is found that the average deviation of the SIFT-TPS displacement from the known solution is less than 1.2 mm.

III.B. Clinical case study

The contour propagation study from pCT to CBCT for the first prostate case is presented in Fig. 7. The top row shows the pCT image with manually outlined contours. The automapped contours overlaid on the CBCT are displayed in the bottom row. For comparison, the manually outlined contours on the CBCT are also plotted in the bottom row. As mentioned in Sec. I, the propagation of rectum wall is often complicated by the fact that the physical one-to-one correspondence may not exist due to the addition or subtraction of some contents within the rectum. Figure 8 exemplifies this and shows that the rectal filling at the time of CBCT acquisition is quite different from that of pCT. As can be intuitively conceived, this image content change could severely

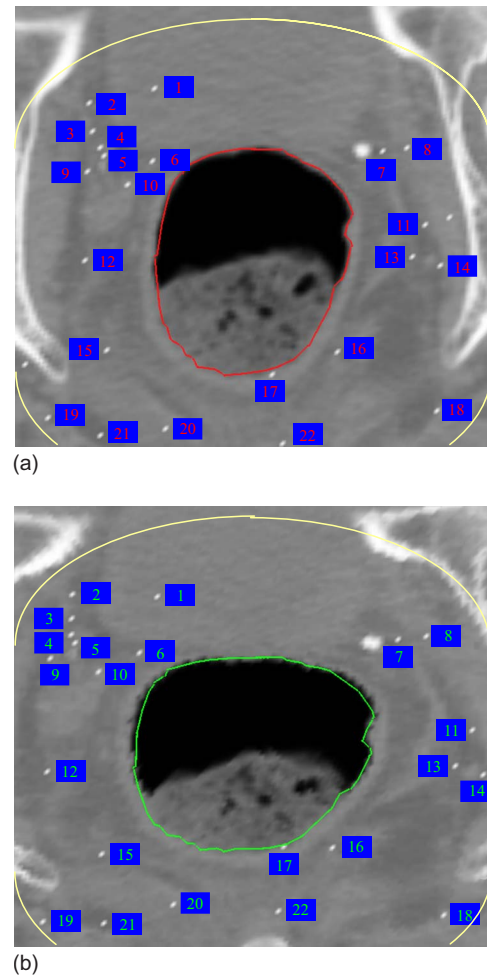


FIG. 5. Control points in the 2D contour mapping.

reduce the performance of a conventional deformable registration.¹⁴ The narrow band approach described in this work circumvents the problem by excluding the rectal volume affected by the rectum/bladder filling. Accuracy was evaluated by comparison with manually outlined contours on the CBCTs.^{15–17} It is clear that the mapped contours closely conform to the rectal wall change. The accordance between the rectal volume extended by the automapped contour and the manually segmented rectal volume was found to be more than 90%.

In practice, rectal volume motion and deformation can cause large uncertainties pertaining to the adequacy of actual dose delivered to the gross tumor volume as well as to the surrounding normal structures. This issue has been a major obstacle in the implementation of IMRT in rectal cancer. In Fig. 8 six axial pCT and CBCT images of a rectal cancer patient acquired in an interval of two weeks are shown. Large target volume motion and deformation are observed from Fig. 8. The rectal volume in the pCT is found to be three times more than that of the rectal volume in the CBCT and thus represents a challenging situation for any deformable model. The rectal contours are manually outlined in the pCT and mapped to the subsequent CBCT using the proposed method. The first and second rows of Fig. 8 show six

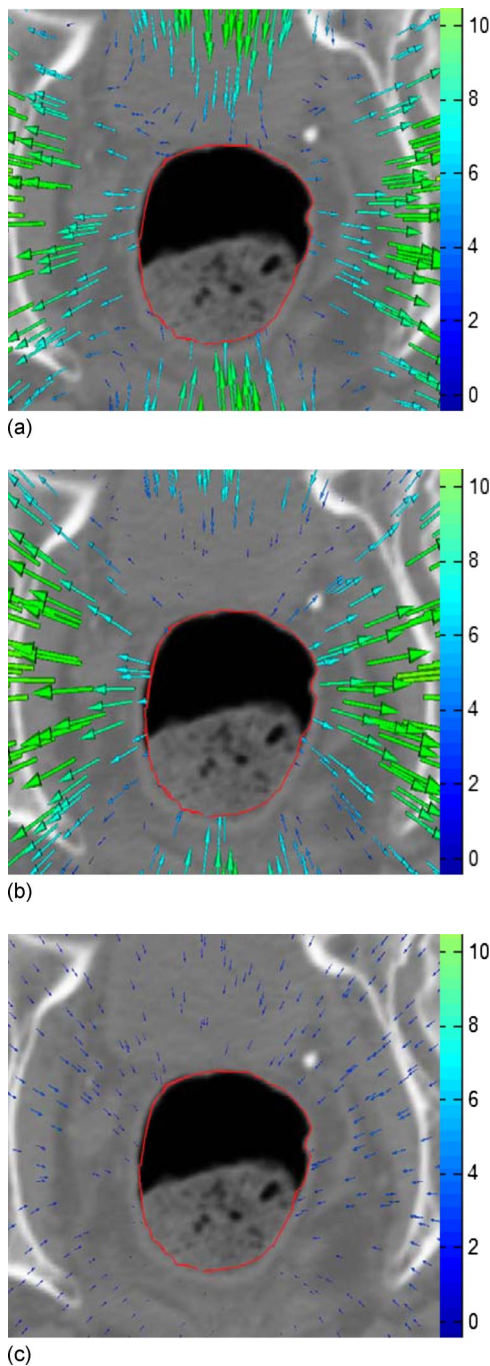


FIG. 6. Displacement fields. (a) TPS-derived displacement field for the 2D digital phantom study; (b) intentionally introduced displacement field; and (c) subtraction of TPS derived and the known displacement fields.

axial slices of the pCT with manually outlined contours. The results of contour propagation from the pCT to the CBCT are shown in the third and fourth rows of Fig. 8. As the same as in Fig. 7 the manually outlined contours on the CBCT are also plotted. The accordance between the rectal volume extended by the automapped contours and the manually segmented rectal volume was found to be more than 95%. The rectal deformations in Fig. 8 are quite large and thus present challenges to any deformable model or contour mapping technique. It is impressive that a simple approach with a

narrow band and SIFT descriptor can capture the main feature of the rectal contour and help to find the correspondence contours in the CBCT images.

To further examine the performance of the proposed technique, the method was also applied to three additional patients (Fig. 9). The automapped contours are plotted together with the manually outlined contour on the CBCT. For comparison, the original contours on pCT are also mapped rigidly to the CBCT. The accordance values between the pCT and CBCT contours, as well as between the automapped and manually segmented CBCT contours for these three patients are listed in Table I. In these cases, the accordance values are increased from around 75% to over 90% after contour mapping. The influence of the κ ratio on the contour propagation is illustrated by the data listed in Table II, where the accordance values for a few different κ ratios for the three patients are shown. The accordance reached its peak value when the κ ratio is between 0.8 and 0.9 for all these three cases. When the κ ratio is lower than 0.8, the accordance decreases with the decrease of the κ ratio because less control points are selected. The accordance also decreases with the increase of the κ ratio for the κ ratio higher than 0.9. The accordance value is stable for κ ratios between 0.8 and 0.9. The data also indicate that the κ ratio is generally organ specific and is insensitive for different patients.

IV. DISCUSSION

In this work, an effective feature-based rectal contour mapping algorithm has been described. An indispensable step toward online or offline adaptive replanning with consideration of the patient's dose delivery history and on-treatment anatomy is the expedite organ segmentation of CBCT images. While this task is, in principle, achievable using deformable registration of the pCT and CBCT images, the accuracy of the registration and therefore the contour mapping, is often adversely affected by the presence of image contents in one image that do not have correspondence in the other image. The propagation of rectum wall is an example of this. For prostate, rectal, or gynecological cancer patients for example, the presence and absence of bowel gas can vary daily. Coregistering an empty rectum without bowel gas to a rectum filled with bowel gas (or vice versa) using any deformable model could be problematic and large errors could occur.

We describe a regional contour propagation algorithm taking into account possible organ deformation and anatomic changes. Because the narrow band contains only the image features outside the rectum, this method is not affected by the rectum filling changes. The use of SIFT descriptor enhances our ability to find the correspondence of the narrow band because of the effective utilization of image intensity and gradient information. In contrast to the conventional intensity-based image registration, which only uses intensity information of the voxels, the feature-based registration extracts information regarding image structure, including shape, texture, etc. Therefore, the feature-based image regis-

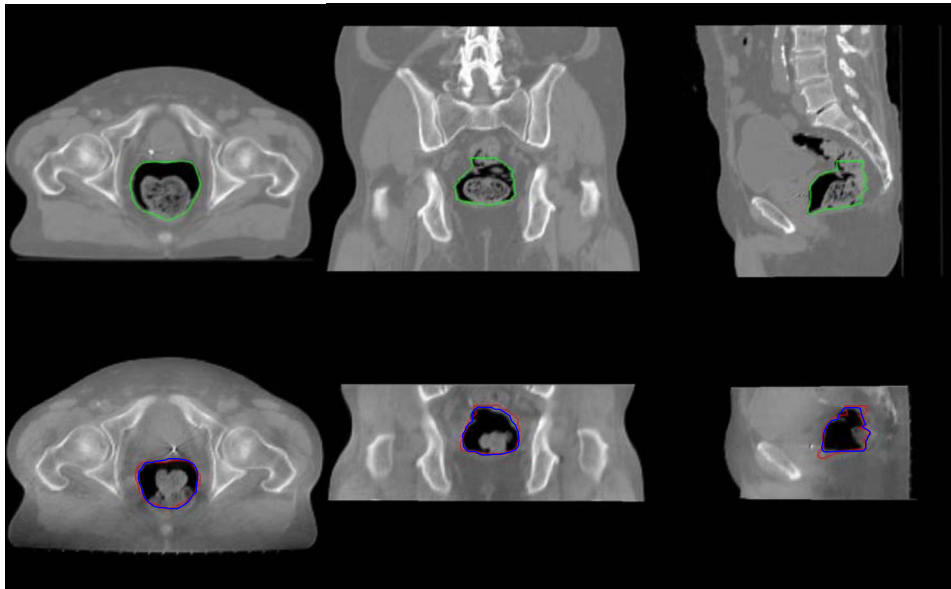


FIG. 7. 3D contour mapping for the rectum of a man with prostate cancer. The top row is the three transactions in the planning CT image, the bottom row is corresponding transactions in the CBCT image. The left column is the axial plane, the middle column is the coronal plane, and the right column is the sagittal plane.

tration is generally more effective in correctly identifying corresponding voxels compared to the intensity-based image registration.

In this study, a bidirectional SIFT descriptor is employed to examine the reliability and robustness of the calculations. The bidirectional mapping further enhances the degree of success of a contour propagation algorithm. It is useful to note that the bidirectional mapping is a necessary (but not

sufficient) test. In a rare but possible situation, the bidirectional mapping may not be able to find an error occurred in the contour mapping process.

Because the iterative procedure in the B-spline is not needed in our method, the calculation speed is at least ten times faster than B-spline registration. Typically the total calculation time of SIFT-TPS mapping with about 1000 SIFT descriptors is less than 2 min. Several parameters influence

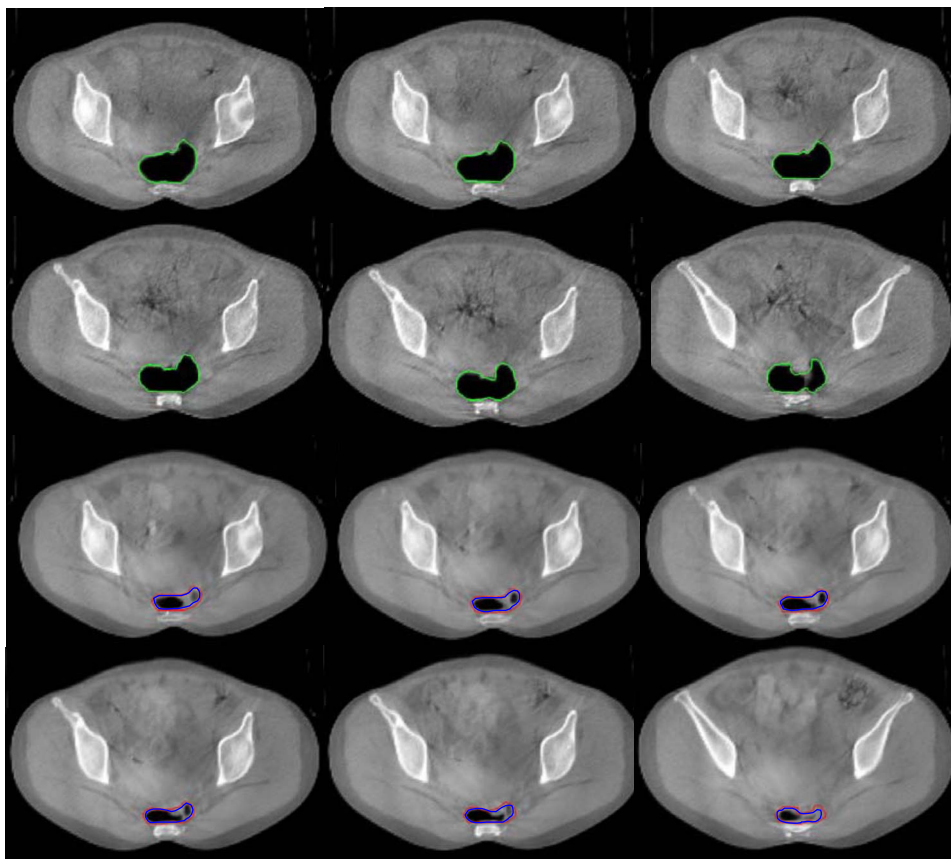


FIG. 8. Rectal contour mapping for a rectal cancer case. The first and second rows show six axial slices in the pCT image. The third and fourth rows are the corresponding slices in the CBCT image.

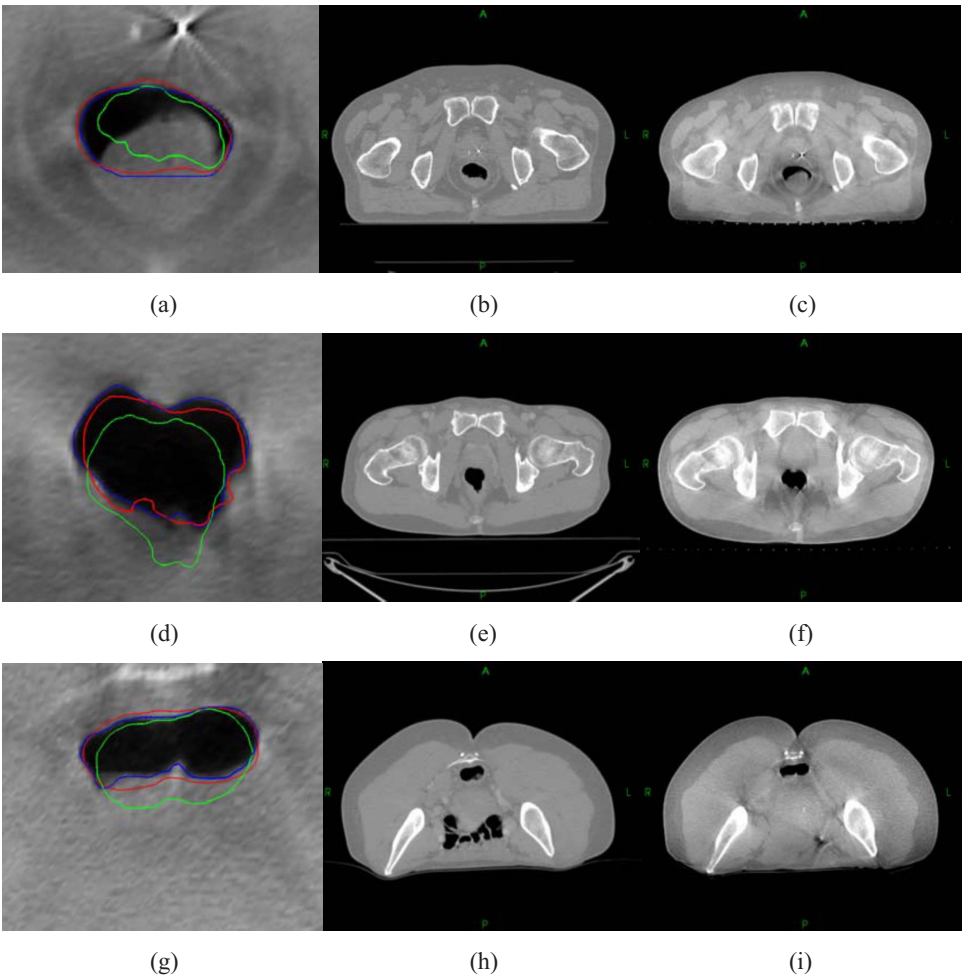


FIG. 9. Automatic contour propagation for three additional patients.

calculation time. For example, larger narrow-band will result in longer calculation times. The number of control points also affects calculation time quite nonlinearly as well. For most cases, 300 control points are enough for accurate contour mapping. Due to the tight clinical timeframes (especially for real-time adaptive schemes), 1 or 2 min calculation time allows the use of the contour mapping tool between acquiring the verification images and delivering the dose fraction for online corrections.¹⁸

In some cases, no corresponding feature is found by SIFT in a certain area close to the rectal wall. For instance, no control point was found in the upper part of Fig. 5. Since no large local deformation was in these regions, the result was all right. However, it would have lead to larger errors in case

of large deformations plus low feature density, which may happen in smooth soft tissues. We will improve it in our future work.

One of the practical concerns is that the relatively low quality of CBCT images may influence the accuracy of image registration and thus the contour mapping. Paquin *et al.* quantitatively studied the influence of different types of noises on deformable registration and found that the accuracy of image registration does not depend on the global noise unless the noise reaches a certain threshold value.¹⁹ Murphy *et al.* also demonstrated that noise levels in cone-beam CTs that might reduce manual contouring accuracy do not reduce image registration and automatic contouring accuracy.²⁰

TABLE I. Accordance values between the pCT and CBCT contours, as well as between the auto-mapped and manually segmented CBCT contours for three patients.

	Patient 1	Patient 2	Patient 3
Accordance values between pCT and CBCT contours (%)	73.7	76.5	76.3
Accordance values between auto-mapped and manually segmented CBCT contours (%)	93.3	91.3	91.4

TABLE II. Accordance values between the auto-mapped and manually segmented CBCT contours for different κ -ratios for three patients.

κ ratio	0.7	0.75	0.8	0.85	0.9	0.95
Patient 1 (%)	81.5	88.4	91.6	92.6	93.3	82.9
Patient 2 (%)	76.3	82.1	91.3	88.9	85.8	79.1
Patient 3 (%)	67.5	86.6	88.3	88.4	91.4	62.6

Deformable model plays an important role in automated contour propagation. Numerous approaches have been developed for different applications. Most popular deformable registration methods for medical images include the thin plate splines (TPS),⁸ B-splines,^{21,22} and finite element method (FEM).²³ TPS is less sensitive to noise because of its global calculation nature.²⁴ It relies on the use of homologous control points in the two input image sets to be coregistered. Control points are manually selected for many TPS applications.^{13,25,26} This may introduce interuser variability and is a major source of error. Malsch *et al.* presented an automatic block matching method,¹⁸ which is similar to the control volume based approach proposed by Schreiber and Xing.²⁷ Kim *et al.* presented an automated TPS, where an arbitrary set of control points is supplied initially and then is iteratively repositioned until the resulting warp optimizes some measure of registration.^{28–30} The convergence of the iterative calculation is slow because each control point influences the transformation in a global fashion. An alternative is to use B-splines. In contrast to TPS, which allows arbitrary configurations of the control points, B-spline requires a regular mesh of control points with uniform spacing. Unlike spline-based registration methods, FEM models the deforming image as an elastic body subject to external forces which drive the deformation and internal forces (stresses) which impose smoothness constraints.^{31,32} FEM may fail to model highly localized deformations, since the deformation energy caused by stress increases proportionally with the strength of the deformation.³³

V. CONCLUSION

Large interfractional patient setup uncertainty and anatomy changes have been reported in numerous studies and are widely recognized as one of the major limiting factors for maximum exploitation of modern radiation therapy techniques such as IMRT and IGRT. The advent of onboard volumetric imaging devices promises to improve the situation by providing valuable 3D (or even possibly four-dimensional) geometric data of the patient in the treatment position and allows for the adaptive modification of treatment plan during a course of treatment.

In this work, an effective feature-based rectal contour mapping algorithm has been described. The method yielded satisfactory mapping for both digital phantom and clinical cases. It is impressive that the algorithm is able to successfully map the contours from pCT to CBCT even for some very challenging cases in which the deformation and/or im-

age content change are dramatic. The two salient features of the described algorithm are: (1) the use of inherent tissue feature for control point selection as *a priori* knowledge for deformable registration; and (2) limiting the ROI to exclude the volume inside the rectum and focusing on the adjacent neighborhood of the rectal contour. The algorithm should be extendable for contour propagation of organs with similar features, such as the bladder and stomach.

ACKNOWLEDGMENTS

This work was supported in part by grants from the Department of Defense (W81XWH-06-1-0235 and W81XWH05-1-0041) and National Cancer Institute (1R01 CA98523 and CA104205).

^{a)}Electronic mail: lei@reyes.stanford.edu

¹L. Xing *et al.*, "Overview of image-guided radiation therapy," *Med. Dosim.* **31**, 91–112 (2006).

²A. de la Zerda, B. Armbruster, and L. Xing, "Formulating adaptive radiation therapy (ART) treatment planning into a closed-loop control framework," *Phys. Med. Biol.* **52**, 4137–4153 (2007).

³P. Lee *et al.*, "Image-guided radiation therapy (RT) for rectal cancer using cone beam CT (CBCT)," *Int. J. Radiat. Oncol., Biol., Phys.* **66**, S276 (2006).

⁴M. Foskey *et al.*, "Large deformation three-dimensional image registration in image-guided radiation therapy," *Phys. Med. Biol.* **50**, 5869–5892 (2005).

⁵S. Gao *et al.*, "A deformable image registration method to handle distended rectums in prostate cancer radiotherapy," *Med. Phys.* **33**, 3304–3312 (2006).

⁶L. Ibanez, W. Schroeder, and L. Ng, "ITK Software Guide," Kitware, Inc., 2003.

⁷W. Schroeder, K. Martin, and B. Lorensen, *The Visualization Toolkit: An Object-Oriented Approach To 3D Graphics*, 3rd ed., 2003.

⁸F. L. Bookstein, "Principal warps: Thin plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 567–585 (1989).

⁹D. G. Lowe, "Object recognition from local scale-invariant features," *Proc. of the International Conference on Computer Vision*, Corfu, 1999.

¹⁰K. Mikołajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.* **27**, 1615–1630 (2005).

¹¹G. Wu, F. Qi, and D. Shen, "A general learning framework for non-rigid image registration," *MIAR*, 2006, pp. 219–227.

¹²W. Lu *et al.*, "Fast free-form deformable registration via calculus of variations," *Phys. Med. Biol.* **49**, 3067–3087 (2004).

¹³J. Lian *et al.*, "Mapping of the prostate in endorectal coil-based MRI/MRSI and CT: A deformable registration and validation study," *Med. Phys.* **31**, 3087–3094 (2004).

¹⁴L. E. Court and L. Dong, "Automatic registration of the prostate for computed-tomography-guided radiotherapy," *Med. Phys.* **30**, 2750–2757 (2003).

¹⁵E. Schreiber, G. T. Chen, and L. Xing, "Image interpolation in 4D CT using a BSpline deformable registration model," *Int. J. Radiat. Oncol., Biol., Phys.* **64**, 1537–1550 (2006).

¹⁶W. Lu *et al.*, "Automatic re-contouring in 4D radiotherapy," *Phys. Med. Biol.* **51**, 1077–1099 (2006).

- ¹⁷T. Zhang *et al.*, "Automatic delineation of on-line head-and-neck computed tomography images: toward on-line adaptive radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **68**, 522–530 (2007).
- ¹⁸U. Malsch *et al.*, "An enhanced block matching algorithm for fast elastic registration in adaptive radiotherapy," *Phys. Med. Biol.* **51**, 4789–4806 (2006).
- ¹⁹D. Paquin, L. Xing, and D. Levy, "Multiscale deformable registration of noisy medical images," *Math. Biosci. Eng.* **5**, 125–144 (2008).
- ²⁰M. J. Murphy *et al.*, "How does CT image noise affect 3D deformable image registration for image-guided radiotherapy planning?," *Med. Phys.* **35**, 1145–1153 (2008).
- ²¹M. Staring, S. Klein, and J. P. W. Pluim, "A rigidity penalty term for nonrigid registration," *Med. Phys.* **34**, 4098–4108 (2007).
- ²²E. Schreibmann *et al.*, "Image interpolation in 4D CT using a B spline deformable registration model," *Med. Phys.* **32**, 1924 (2005).
- ²³R. Alterovitz *et al.*, "Registration of MR prostate images with biomechanical modeling and nonlinear parameter estimation," *Med. Phys.* **33**, 446–454 (2006).
- ²⁴S. Roberts and L. Stals, "Discrete thin plate spline smoothing in 3D," *Annal. Chim.* **45**, C646–C659 (2004).
- ²⁵B. Fei, C. Kemper, and D. L. Wilson, "A comparative study of warping and rigid body registration for the prostate and pelvic MR volumes," *Comput. Med. Imaging Graph.* **4**, 267–281 (2003).
- ²⁶M. M. Coselmon *et al.*, "Mutual information based CT registration of the lung at exhale and inhale breathing states using thin-plate splines," *Med. Phys.* **31**, 2942–2948 (2004).
- ²⁷E. Schreibmann and L. Xing, "Image registration with auto-mapped control volumes," *Med. Phys.* **33**, 1165–1179 (2006).
- ²⁸B. Kim *et al.*, "Mutual information for automated unwarping of rat brain autoradiographs," *Neuroimage* **5**, 31–40 (1997).
- ²⁹C. R. Meyer *et al.*, "Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations," *Med. Image Anal.* **1**, 195–206 (1997).
- ³⁰K. M. Brock *et al.*, "Automated generation of a four-dimensional model of the liver using warping and mutual information," *Med. Phys.* **30**, 1128–1133 (2003).
- ³¹L. Xing, J. Siebers, and P. Keall, "Computational challenges for image-guided radiation therapy: Framework and current research," *Semin. Radiat. Oncol.* **17**, 245–257 (2007).
- ³²H. Lester and S. R. Arridge, "A survey of hierarchical non-linear medical image registration," *Pattern Recogn.* **32**, 129–149 (1999).
- ³³J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, *Medical Image Registration*. (CRC Press, Boca Raton, 1999).
- ³⁴ITK, <http://www.itk.org>.
- ³⁵VTK, <http://public.kitware.com/VTK/>.

Dose reduction for kilovoltage cone-beam computed tomography in radiation therapy

Jing Wang¹, Tianfang Li², Zhengrong Liang³ and Lei Xing¹

¹ Department of Radiation Oncology, Stanford University School of Medicine, Clinical Cancer Center, 875 Blake Wilbur Drive, Rm CC-G204, Stanford, CA 94305-5847, USA

² Department of Radiation Oncology, University of Pittsburgh Medical Center, Pittsburgh, PA 15901, USA

³ Department of Radiology, State University of New York, Stony Brook, NY 11794, USA

E-mail: lei@reyes.stanford.edu

Received 19 December 2007, in final form 12 March 2008

Published 6 May 2008

Online at stacks.iop.org/PMB/53/2897

Abstract

Kilovoltage cone-beam computed tomography (kV-CBCT) has shown potentials to improve the accuracy of a patient setup in radiotherapy. However, daily and repeated use of CBCT will deliver high extra radiation doses to patients. One way to reduce the patient dose is to lower mAs when acquiring projection data. This, however, degrades the quality of low mAs CBCT images dramatically due to excessive noises. In this work, we aim to improve the CBCT image quality from low mAs scans. Based on the measured noise properties of the sinogram, a penalized weighted least-squares (PWLS) objective function was constructed, and the ideal sinogram was then estimated by minimizing the PWLS objection function. To preserve edge information in the projection data, an anisotropic penalty term was designed using the intensity difference between neighboring pixels. The effectiveness of the presented algorithm was demonstrated by two experimental phantom studies. Noise in the reconstructed CBCT image acquired with a low mAs protocol was greatly suppressed after the proposed sinogram domain image processing, without noticeable sacrifice of the spatial resolution.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Integration of the kilovoltage cone-beam computed tomography (kV-CBCT) with a linear accelerator makes it possible to acquire a high-resolution volumetric image of a patient at a treatment position. There is growing interest in using on-board kV-CBCT for a patient treatment position setup and dose reconstruction in radiotherapy (Xing *et al* 2006, Yang *et al* 2007, Lee *et al* 2008). However, the repeated use of kV-CBCT during the course of a treatment

has raised concerns of an extra radiation dose delivered to patients (Brenner and Hall 2007, Islam *et al* 2006, Wen *et al* 2007). It has been reported (Wen *et al* 2007) that the dose delivered from Varian's kV-CBCT system with current clinical protocols is more than 3 cGy for central tissue and about 5 cGy for most of the peripheral tissues during an IMRT (intensity-modulated radiation therapy) treatment course for prostate cancer. The extra radiation exposure to normal tissue during kV-CBCT will significantly increase the probability of stochastic risk of inducing cancer and genetic defects. Based on the ALARA (as low as reasonably achievable) principle, the unwanted kV-CBCT dose should be minimized to fully realize its advantages of precise target localization during radiotherapy (Murphy *et al* 2007).

One way to reduce the radiation dose delivered to patients during the kV-CBCT procedure is to acquire CT projection data with a lower mAs level (can be realized by reducing the tube current or pulse time). However, the image quality of the projection image and the reconstructed CBCT image will be degraded due to excessive quantum noise as a result of a low mAs protocol. Conventionally, noise in CT is suppressed by using a low-pass filter to attenuate the high-frequency component of the projection data during reconstruction. The high-frequency component contains information of both noise and edges, where a simple low-pass filter cannot differentiate edge information from noise. Therefore, noise reduction using a low-pass filter will result in loss of edges, which is not desirable for CT imaging. Several edge-preserving filters (Hsieh 1998, Kachelriess *et al* 2001, Zhong *et al* 2004) have been proposed to reduce noise in CT images based on local characteristics of the projection data elements. More recently, statistics-based image domain (Li *et al* 2005a) and sinogram domain restoration algorithms (Li *et al* 2004, La Riviere 2005, La Riviere and Billmire 2005, Wang *et al* 2006) have shown advantages in noise reduction and edge preservation for low-dose fan-beam CT. In the meantime, noise properties of CT projection data have been under investigation (Li *et al* 2004, Whiting *et al* 2006) and the noise model of the sinogram data in Radon space (i.e. line integrals) has been validated by experimental studies (Wang *et al* 2008). In this work, we aim to improve the low-dose CBCT image quality by reducing noise in the CBCT sinogram before image reconstruction. The noise reduction algorithm incorporates the noise modeling of the CT sinogram data in Radon space (line integrals) to construct a penalized weighted least-squares (PWLS) objective function (Fessler 1994, Sukovic and Clinthorne 2000). The ideal solution of the line integrals is then estimated by minimizing the PWLS objective function. The weighted least square is based on the first and second moments of the noise in the sinogram data and an anisotropic penalty is designed to preserve the edges in the sinogram. CBCT images are reconstructed by using the Feldkamp–Davis–Kress (FDK) (Feldkamp *et al* 1984) algorithm after all sinogram images are processed by the PWLS criterion sequentially. The effectiveness of the PWLS-based noise reduction algorithm is demonstrated by two experimental phantom studies.

2. Methods and materials

2.1. CBCT sinogram smoothing

Ideally, the line integral of attenuation coefficients can be calculated by

$$p_i = \ln \frac{N_{i0}}{N_i}, \quad (1)$$

where N_{i0} and N_i are the incident photon number and detected photon number at the detector bin i respectively. For ease of presentation, we refer the measurement as a photon number. In a real x-ray CBCT system, the measured signal is total energy deposit on the flat-panel detector.

In the following of this paper, we refer the value of p_i as the sinogram datum at the detector bin i . Mathematically, the PWLS cost function in the sinogram domain can be written as

$$\Phi(p) = (\hat{y} - \hat{p})^T \Sigma^{-1} (\hat{y} - \hat{p}) + \beta R(p). \quad (2)$$

The first term in equation (2) is a weighted least-squares criterion, where \hat{y} is the vector of the measured sinogram data and \hat{p} is the vector of the ideal sinogram data to be estimated. The symbol T denotes the transpose operator. The matrix Σ is a diagonal matrix and its i th element is the variance of sinogram data at the detector bin i . The second term in equation (2) is a smoothness penalty or *a priori* constraint, where β is the smoothing parameter which controls the degree of agreement between the estimated and the measured data.

The element of the diagonal matrix Σ is the variance of the corresponding sinogram datum, and it determines the contribution of each sinogram datum to the cost function. Based on the sinogram noise modeling in Li *et al* (2004) and Wang *et al* (2008), the variance of the sinogram datum can be estimated by

$$\sigma_i^2 = \exp(p_i) / N_{i0}. \quad (3)$$

For a fixed incident photon number N_{i0} , a sinogram datum with a larger value will have a larger variance and therefore less contribution to the cost function since the weight of that measured datum is $1/\sigma_i^2$ as defined in equation (2). This can be understood by the following observation. A larger sinogram datum value p_i at the detector bin i indicates less x-ray photons being detected, i.e. smaller N_i in equation (1), or more photons being attenuated along the projection path i . A detector bin receiving less photons will be associated with a smaller signal-to-noise ratio (SNR) based on the Poisson noise nature of the detected x-ray photons. Therefore, the weighted least-squares criterion reflects the above observation that the measured datum with a lower SNR will contribute less for estimation of its ideal sinogram datum.

To calculate the sinogram datum variance at the detector bin i via equation (3), we need to estimate the incident photon number N_{i0} for calculation of the sinogram variance. The incident photon number is mainly determined by the protocols of tube current and the duration of x-ray pulse (i.e. mAs). Ideally, the incident x-ray flux from the tube would be calibrated as uniform as possible across a field of view (FOV), i.e. N_{i0} is a constant for all the detector bins. In reality, the x-ray flux is modulated to consider the concavity shape of the human body by the bow-tie attenuating filter prior to arrival at the patients. Therefore, the incident photon number will not be a constant across the FOV. To estimate the incident intensity over the FOV at a specific mAs level, we performed the air scan and then averaged the projection image over all projection view angles. Figure 1 shows the incident x-ray intensity with the tube current 80 mA and duration of pulse 10 ms. The incident x-ray intensity can then be used for estimation of the sinogram data variance $\{\sigma_i^2\}$.

The penalty term in equation (2) is a prior or smoothing constraint, which encourages the equivalence between neighboring data elements. In Li *et al* (2004) and Wang *et al* (2006), a penalty of a quadratic form with equal weights for all neighbors has been used for sinogram smoothing of fan-beam CT:

$$R(p) = \sum_n w_{in} (p_i - p_n)^2, \quad (4)$$

where n represents four nearest neighbors around pixel i and w_{in} is the weight for neighbor n . With an equal weight for the four nearest neighbors, these neighbors play an equivalent role in constraining the solution. As such, it provides a uniform regularization without considering details of intensity variation and possibly the presence of edges in the sinogram image. To preserve the edge information in the sinogram image of CBCT, we propose to use

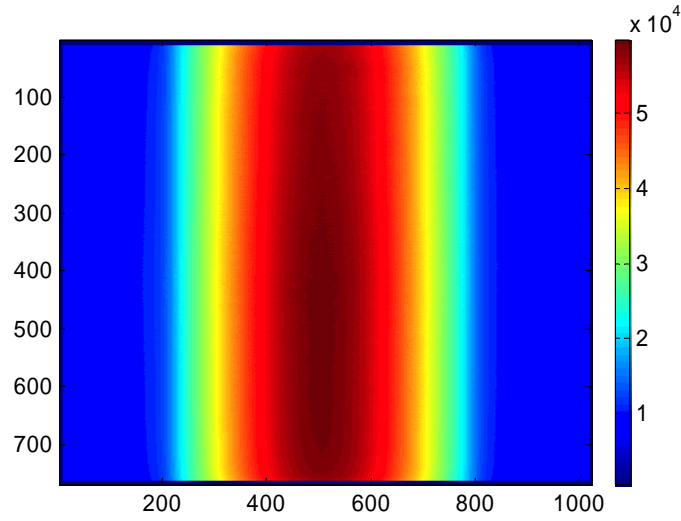


Figure 1. Incident x-ray intensities across the field of view with 80 mA tube current and 10 ms pulse time. Relative intensity is mainly caused by the bow-tie filter.

anisotropic weights for different neighbors in the sinogram image. The weight of the neighbor is determined by the magnitude of difference between neighbors and the concerned pixel. For a larger difference between the neighbor and the pixel, the coupling between them should be weaker and the weight w_{in} should be smaller. This form of weight is chosen the same as the conducting coefficient in the well-known anisotropic diffusion filter (Perona and Malik 1990):

$$w_{in} = \exp \left[- \left(\frac{p_i - p_n}{\delta} \right)^2 \right], \quad (5)$$

where the gradient determines the strength of the diffusion during each iteration and the parameter δ was chosen as 90% of histogram of the gradient magnitude of the sinogram to be processed (Perona and Malik 1990).

Minimization of the objective function 2 can be performed efficiently by the iterative Gauss–Seidel updating strategy. The updating formula for the solution of \hat{p} is given by

$$p_i^{(k+1)} = \frac{y_i + \beta \sigma_i^2 \left(\sum_{n \in N_i^1} w_{in} p_n^{(k+1)} + \sum_{n \in N_i^2} w_{in} p_n^{(k)} \right)}{1 + \beta \sigma_i^2 \sum_{n \in N_i} w_{in}}, \quad (6)$$

where the index k represents the iterative number, N_i^1 denotes those two nearest neighbors of i whose index is smaller than i , N_i^2 denotes those two nearest neighbors of i whose index is larger than i and N_i denotes these four nearest neighbors of pixel i in the sinogram image. The initial of \hat{p} is given by the measured data \hat{y} .

2.2. On-board kV-CBCT

The cone-beam CT projection data were acquired by ExactArms (kV source/detector arms) of a Trilogy(tm) treatment system (Varian Medical Systems, Palo Alto, CA). The number of projections for a full 360° rotation is around 634. The dimension of each acquired projection image is 397 mm × 298 mm, containing 1024 × 768 pixels. The system has a FOV of 25 cm × 25 cm (full-fan mode) in the transverse plane and 17 cm in the longitudinal direction, which

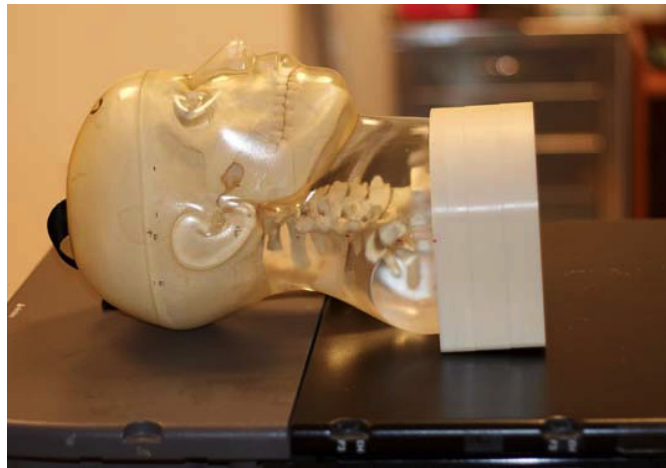


Figure 2. Illustration of the anthropomorphic head phantom used for evaluation of the PWLS algorithm.

can be increased to $45\text{ cm} \times 45\text{ cm}$ in the transverse plane by shifting the detector laterally (half-fan mode).

Two phantoms were used to evaluate the performance of the proposed PWLS algorithm in this study. The first phantom is a commercial calibration phantom CatPhan[®] 600 (The Phantom Laboratory, Inc., Salem, NY). Details about the CatPhan[®] 600 phantom can be found in Li *et al* (2005a). The second one is an anthropomorphic head phantom (see figure 2). For each phantom, the x-ray tube current was set at 10 mA (low dose) and 80 mA (high dose) during acquisition of CBCT projection images. At both mA levels, the duration of the x-ray pulse at each projection view was 10 ms. The tube voltage was set to 125 kVp during all data acquisitions. After each sinogram acquired with the low-mAs protocol was processed by the PWLS algorithm described above, the CBCT image was reconstructed by the FDK algorithm. The voxel size in the reconstructed image is $0.5 \times 0.5 \times 0.5\text{ mm}^3$.

3. Results

3.1. CatPhan[®] 600 phantom

We first tested the proposed algorithm on the CatPhan[®] 600 phantom. Several representative slices of the reconstructed CBCT are shown in figures 3, 4 and 6. In each of these figures, (a) is the FDK reconstructed image from the projection data acquired with 10 mA tube current, (b) is the FDK reconstructed image from the sinogram processed by the proposed PWLS sinogram smoothing algorithm and (c) is the FDK reconstructed image from the sinogram obtained with 80 mA tube current.

Figure 3 shows that one slice of image contains a point-like object, which mimics a fiducial marker. In figure 3(a), the point source is difficult to be observed. After the sinogram was processed by the PWLS algorithm, the reconstructed image (figure 3(b)) is very similar to that obtained with a high mA protocol (figure 3(c)). The point source was well recovered and easy to be detected.

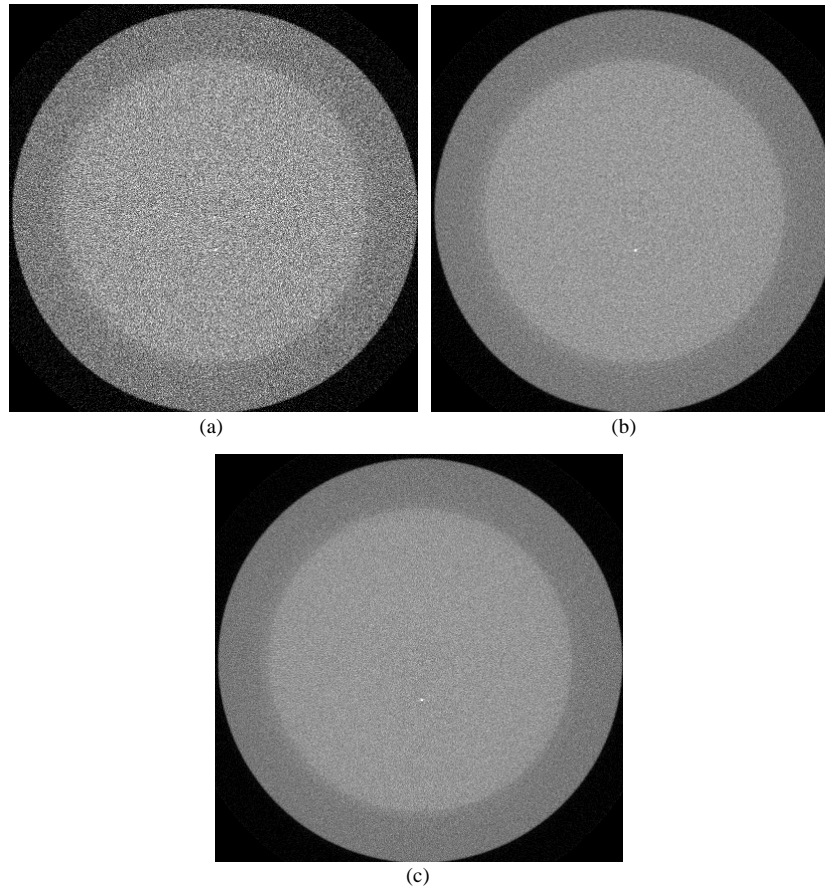


Figure 3. One slice of the FDK reconstructed image of the CatPhan® 600 phantom containing a point-like object: (a) from projection images acquired with 10 mA tube current, (b) after the sinogram acquired with 10 mA tube current is processed by the PWLS algorithm and (c) from projection images acquired with 80 mA tube current.

Figure 4 shows that one slice of image contains several strips with different sizes and contrasts, which can be used to study the edge information in the reconstructed images. The CT image reconstructed from the PWLS-processed sinogram is comparable to that obtained with the 80 mA protocol in terms of detectability of the strips; see ROI2 in figure 4(c). To show the difference between figures 4(a), (b) and (c), in figure 5 we plotted horizontal profiles along the central strips (see ROI1 in figure 4(c)). It can be observed that the edges are well preserved (compare profiles through figures 3(b) and (c)), while noise is effectively suppressed (compare profiles through figures 3(a) and (b)).

To further quantitatively evaluate the effectiveness of the PWLS sinogram smoothing algorithm, we calculated the contrast-to-noise ratio (CNR) at different regions of interest (ROIs) in the images shown in figure 6. The CNR is defined as

$$\text{CNR} = \frac{|\mu_s - \mu_b|}{\sqrt{\sigma_s^2 + \sigma_b^2}}, \quad (7)$$

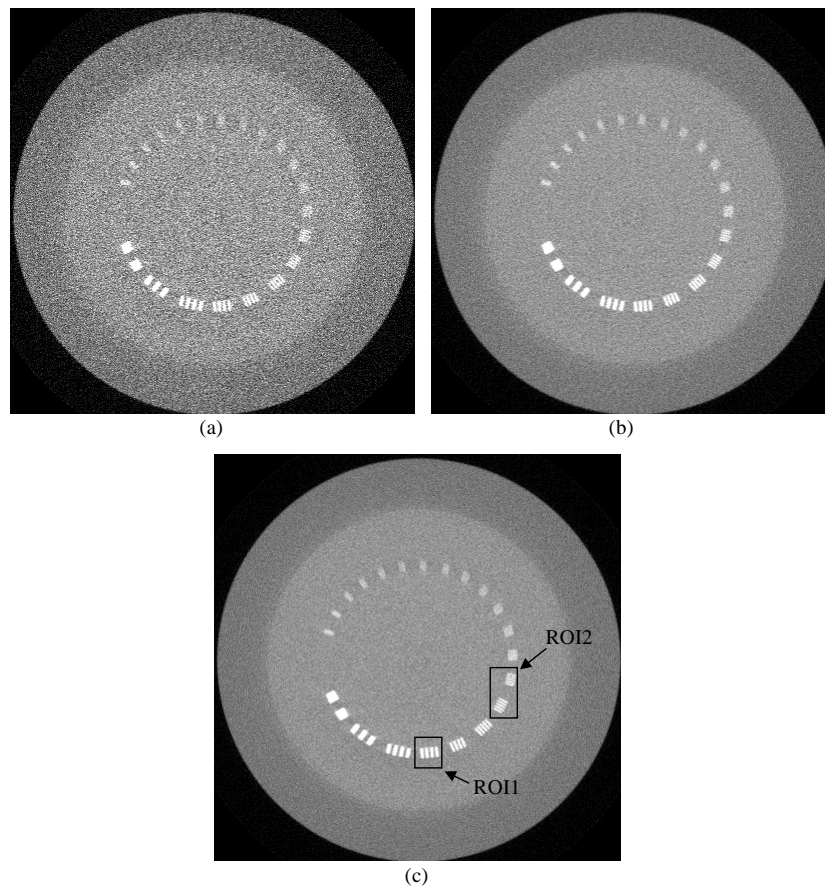


Figure 4. One slice of the FDK reconstructed image of the CatPhan[®] 600 phantom containing several strips: (a) from projection images acquired with 10 mA tube current, (b) after the sinogram acquired with 10 mA tube current is processed by the PWLS algorithm and (c) from projection images acquired with 80 mA tube current.

Table 1. CNRs of four ROIs in figure 5.

	ROI1	ROI2	ROI3	ROI4	ROI5
80 mA	1.83	7.31	4.75	1.51	0.89
10 mA	0.82	2.70	1.68	0.49	0.36
PWLS 10 mA $\beta = 0.05$	1.92	6.88	4.75	1.33	0.85

where μ_s is the mean value of the signal and μ_b is the mean value of the background. Five circular objects (indicated by arrows in figure 6) with different intensities were chosen to calculate CNRs. Table 1 lists the CNRs of these five ROIs. After a 10 mA sinogram was processed by the PWLS algorithm, the CNR in the reconstructed image improved significantly. It can be observed that the CNR of a PWLS-processed 10 mA image is comparable to that of the image acquired with the 80 mA protocol.

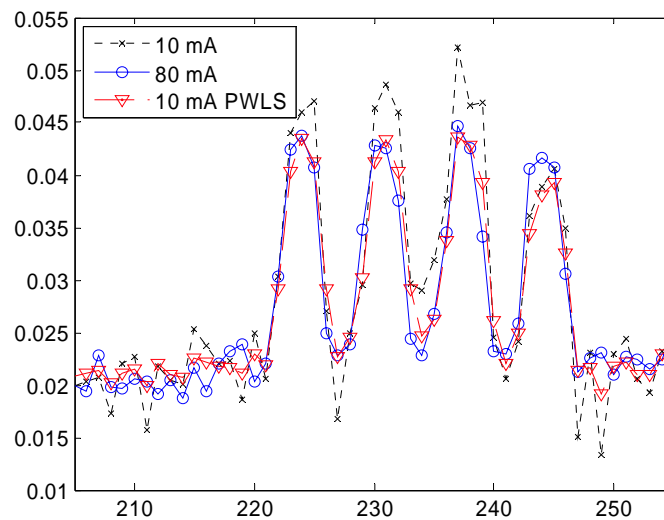


Figure 5. Profiles through the central strips in figure 4 (indicated by ROI1).

3.2. Anthropomorphic head phantom

Results of the anthropomorphic head phantom are shown in figure 7. Figure 7(a) shows one slice of the reconstructed images from projection data acquired with the 10 mA protocol. Figure 7(c) shows the reconstructed image from the PWLS-processed 10 mA sinogram. Figure 7(d) shows the same slice of the image reconstructed from the sinogram obtained with 80 mA. It can be observed that noise in 10 mA CT images is efficiently suppressed after the sinogram is processed by the PWLS algorithm. The processed low-dose CT (10 mA) image is very similar to its corresponding high-dose image (80 mA) by visual judgment. Standard deviation of the noise in a uniform ROI (as indicated by an arrow in figure 7(d)) is 2.8×10^{-3} in a low-dose (10 mA) image and 0.951×10^{-3} in its corresponding high-dose image. After the low-dose sinogram is processed by the PWLS algorithm with a smoothing parameter $\beta = 0.05$, the standard deviation of the same ROI is 0.955×10^{-3} , which is fairly close to the noise level of the 80 mA image.

To further illustrate how the edge information is affected by the PWLS sinogram smoothing, in figure 7(e) we show the difference image between figures 7(a) and (c). In the difference image, random noise is dominant and no edge or structure can be observed. This indicates that the edge information is well preserved in the PWLS-processed images.

4. Discussion

Generally, noise reduction for CT imaging can be performed in three spaces: projection data (either before or after logarithmic transform), filtered projection data (before backprojection operation during reconstruction) and reconstructed CT images. During filtering and backprojection operation, the noise properties will change significantly. Then noise modeling, such as distribution of noise and variance of noise, is difficult in filtered projection data and reconstructed image. Therefore, in this work we chose to work on the log-transformed data to fully utilize the noise model of the projection data in the Radon space (Li *et al* 2004, Wang *et al* 2008).

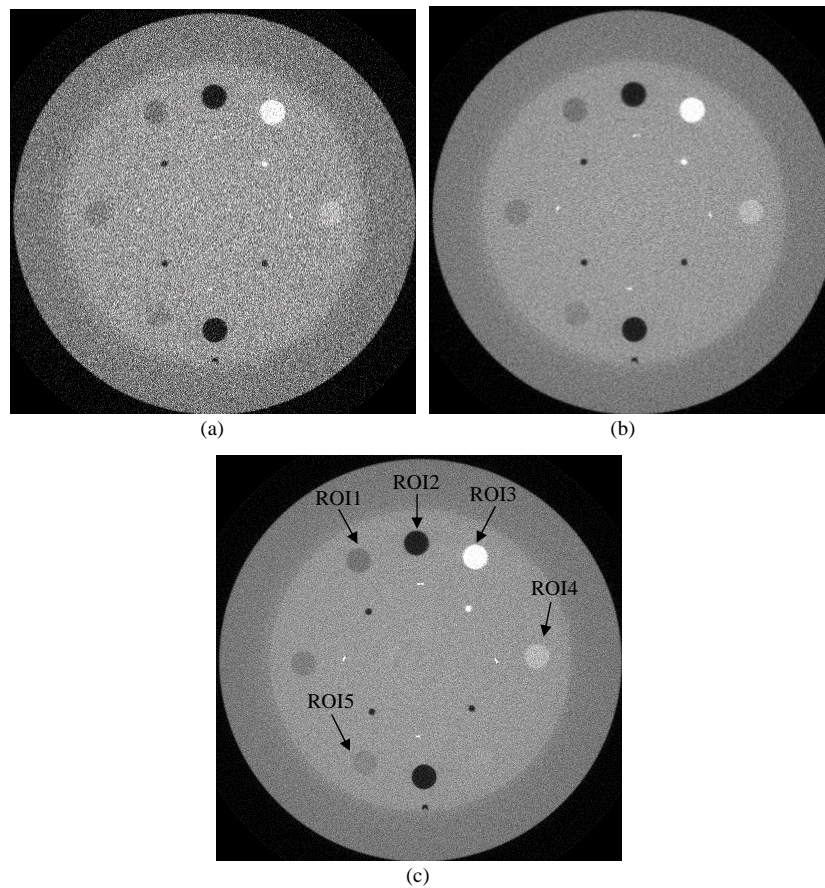


Figure 6. One slice of the FDK reconstructed image of the CatPhan[®] 600 phantom containing several circular objects with different intensities: (a) from projection images acquired with 10 mA tube current, (b) after the sinogram acquired with 10 mA tube current is processed by the PWLS algorithm and (c) from projection images acquired with 80 mA tube current.

Accurate noise modeling of measurement is fundamentally important in statistics-based image processing algorithms. Meanwhile, the regularization term also plays an important role in the performance of the algorithm. In CT sinogram processing, a commonly used regularization takes a quadratic form with equal weights for neighbors of an equal distance (La Riviere 2005, La Riviere and Billmire 2005, Li *et al* 2004, Wang *et al* 2006). Such a quadratic penalty simply encourages the equivalence between neighbors without considering discontinuities in the image and may lead to over-smoothing around sharp edges or boundaries. In the presented algorithm, we proposed an anisotropic penalty to consider the difference among neighbors. The idea was inspired by the well-known anisotropic diffusion filter (Perona and Malik 1990), in which the gradient controls the strength of diffusion among neighbors. The coupling between neighbors should be smaller if the absolute value of difference between them is smaller and this kind of neighbors should contribute less to the solution of the concerned pixel (see equation (6)). There are many choices that satisfy this behavior of weighting. In this work, the form of the anisotropic weight was chosen the same as the conduction coefficients

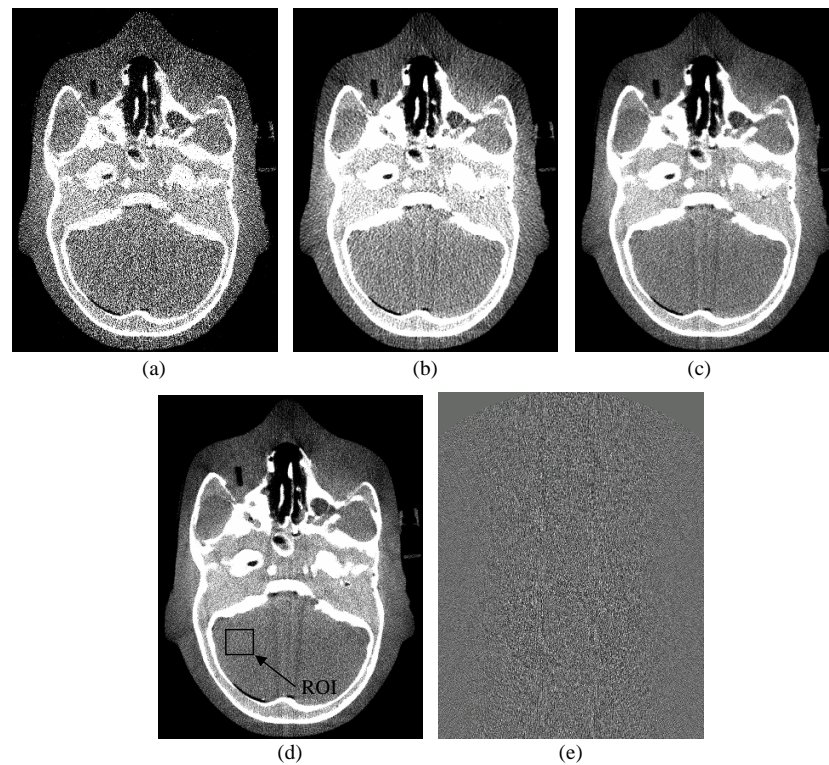


Figure 7. One slice of FDK reconstructed image of the anthropomorphic head phantom: (a) from projection images acquired with 10 mA tube current, (b) using a low-pass Hanning filter with cutoff 80% Nyquist frequency, (c) after the sinogram acquired with 10 mA tube current is processed by the PWLS algorithm, (d) from projection images acquired with 80 mA tube current and (e) difference image between (d) and (c).

in the anisotropic diffusion filter (Perona and Malik 1990). By such a choice, the anisotropic quadratic form penalty discourages the equivalence between neighbors if the gradient between them is large, and the edges or boundaries in the image will be better preserved. This effect is similar to that of anisotropic coefficients in the diffusion filter.

In the presented method, the reconstruction of CT images was performed by an analytical FDK algorithm for its speed and accuracy. During the FDK reconstruction process, noise can also be suppressed by using a low-pass filter. It has been reported (Li *et al* 2004, La Riviere 2005) that a statistics-based sinogram smoothing algorithm plus FBP reconstruction is superior to conventional low-pass filters for noise suppression of 2D fan-beam CT. In this work, we also reconstructed the CT image of the anthropomorphic head phantom using a Hanning filter with a cutoff at 80% Nyquist frequency, see figure 7(b). It can be observed that the image reconstructed from the PWLS-processed sinogram is superior to the result of the Hanning filter in terms of noise suppression and structure preservation.

Similar to the cutoff frequency in the conventional low-pass filter during reconstruction, there is also a free parameter β in the presented method which controls the trade-off of the noise level and the structure preservation in reconstructed images. In this work, the choice of β is determined by the visual judgment. The optimal choice of the parameter β can be determined by more sophisticated ways such as the received operating characteristic (ROC)

study. Nevertheless, the parameter β can be chosen according to the noise level of the sinogram because from equation (6), the solution for the ideal sinogram, we can see that the parameter β and variance σ_i^2 are always coupled together. The noise level of the projection data is mainly determined by two factors: incident photon number and line integrals. As such, the parameter β could be optimized at a certain mAs level and treatment site for patients of a similar size. In this work, however, the parameter is chosen empirically, which is justifiable when the dependence of the parameter on the noise level is weak.

The method presented in this paper is based on the noise properties of the sinogram, and the smoothing constraint or penalty is applied to the sinogram domain. Based on the same noise model, the smoothing constraint can also be applied to the CT image domain, and the statistical iterative reconstruction (SIR) algorithm can be used to obtain the attenuation coefficient map by minimizing the objective function. The SIR-based algorithms showed some advantages over the conventional filtered backprojection method for multi-slice helical CT (Thibault *et al* 2007). However, an obstacle for practical use of SIR is the computation burden, especially for large volume datasets of CBCT. It takes more than 10 h to reconstruct the typical volume of multi-slice helical CT using SIR (Thibault *et al* 2007). It takes only 3 s for the presented sinogram smoothing method to process one projection image using a PC with 3.0 GHz CPU. Parallel computing can speed up both SIR and sinogram-based algorithms significantly using the cell broadband engine (Knaup *et al* 2006) and PC cluster (Li *et al* 2005b). It is possible to achieve clinically acceptable time for the presented sinogram smoothing algorithm through parallel computation. It is an interesting research topic to quantitatively compare the performance of the SIR-based CBCT reconstruction algorithm and the statistics-based sinogram smoothing method.

When CBCT is used for patient setup and target localization during radiotherapy, some extra information may be taken into account for dose and noise reduction. For example, a complete CT volume dataset with high clarity used for treatment planning is usually available before the treatment. This will provide strong *a priori* information of the patient before each CBCT scan. Prior information of planning 3D CT has been proved useful to improve the image quality of 4D CBCT (Li *et al* 2007). It is expected that the radiation dose of CBCT used for radiotherapy can be further reduced by incorporating the planning CT information into the image restoration or reconstruction algorithms.

In the report of the AAPM task group 75 (Murphy *et al* 2007), several dose reduction strategies for image-guided radiotherapy were discussed. For CBCT, dose reduction can be achieved by narrowing field of view to avoid delivering radiation to an unnecessary region of the patient (Murphy *et al* 2007). Compared with these hardware-based approaches, software approaches (such as the one proposed in this paper) provide a more cost-effective means for dose reduction of CBCT. In addition to the statistics-based reconstruction and restoration algorithms, advanced analytical CBCT reconstruction algorithms (Leng *et al* 2007, Zhuang *et al* 2006, Zou and Pan 2004, Zou *et al* 2005) may further improve the low-dose CBCT image quality.

5. Conclusion

A PWLS algorithm with non-uniform weights was proposed to reduce noise in low-dose onboard CBCT. In this method, the sinogram was first processed according to the PWLS criterion. The weight for each measurement was chosen as sinogram datum variance, where variance can be estimated accurately according to the sinogram noise model. To preserve edge information during noise reduction, we proposed an anisotropic quadratic form penalty. The quadratic form penalty encourages equivalence between neighbors and the anisotropic

penalty provides the mechanism to control the influence of different neighbors according to its corresponding gradient. The effectiveness of the proposed method is demonstrated by two experimental phantom studies. The quality of the 10 mA CT image after its sinogram processed by the PWLS algorithm is comparable to the image obtained with the 80 mA protocol. These experimental results indicate that it is possible to reduce the CBCT radiation dose by a factor of 1/8 without loss of useful information for radiotherapy.

Acknowledgment

This work was supported in part by grants from the Department of Defense (PC040282 and PC073592) and National Cancer Institute (1R01 CA98523 and CA104205). Dr Liang was supported by the National Cancer Institute under grants CA082402 and CA120917.

References

- Brenner D and Hall E 2007 Current concepts—computed tomography—an increasing source of radiation exposure *New Engl. J. Med.* **357** 2277–84
- Feldkamp L, Davis L and Kress J 1984 Practical cone-beam algorithm *J. Opt. Soc. Am. A* **1** 612–9
- Fessler J 1994 Penalized weighted least-squares image-reconstruction for positron emission tomography *IEEE Trans. Med. Imaging* **13** 290–300
- Hsieh J 1998 Adaptive streak artifact reduction in computed tomography resulting from excessive x-ray photon noise *Med. Phys.* **25** 2139–47
- Islam M, Purdie T, Norrlinger B, Alasti H, Moseley D, Sharpe M, Siewerdsen J and Jaffray D 2006 Patient dose from kilovoltage cone beam computed tomography imaging in radiation therapy *Med. Phys.* **33** 1573–82
- Kachelriess M, Watzke O and Kalender W 2001 Generalized multi-dimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam CT *Med. Phys.* **28** 475–90
- Knaup M, Kalender W A and Kachelrieß M 2006 Statistical cone-beam CT image reconstruction using the cell broadband engine *IEEE Nucl. Sci. Symp. Conf. Record* **5** 2837–40
- La Riviere P 2005 Penalized-likelihood sinogram smoothing for low-dose CT *Med. Phys.* **32** 1676–83
- La Riviere P and Billmire D 2005 Reduction of noise-induced streak artifacts in X-ray computed tomography through spline-based penalized-likelihood sinogram smoothing *IEEE Trans. Med. Imaging* **24** 105–11
- Lee L, Le Q and Xing L 2008 Retrospective IMRT dose reconstruction based on cone-beam CT and MLC log-file *Int. J. Radiat. Oncol. Biol. Phys.* **70** 634–44
- Leng S, Zhuang T, Nett B and Chen G 2007 Helical cone-beam computed tomography image reconstruction algorithm for a tilted gantry with N-PI data acquisition *Opt. Eng.* **46** 15004
- Li T, Koong A and Xing L 2007 Enhanced 4D cone-beam CT with inter-phase motion model *Med. Phys.* **34** 3688–95
- Li T, Li X, Wang J, Wen J, Lu H, Hsieh J and Liang Z 2004 Nonlinear sinogram smoothing for low-dose X-ray CT *IEEE Trans. Nucl. Sci.* **51** 2505–13
- Li T, Schreibmann E, Thorndyke B, Tillman G, Boyer A, Koong A, Goodman K and Xing L 2005a Radiation dose reduction in four-dimensional computed tomography *Med. Phys.* **32** 3650–60
- Li X, Ni J and Wang G 2005b Parallel iterative cone-beam CT image reconstruction on a PC cluster *J. X-Ray Sci. Tech.* **13** 63–72
- Murphy M *et al* 2007 The management of imaging dose during image-guided radiotherapy: report of the AAPM Task Group 75 *Med. Phys.* **34** 4041–63
- Perona P and Malik J 1990 Scale-space and edge-detection using anisotropic diffusion *IEEE Trans. Pattern Anal.* **12** 629–39
- Sukovic P and Clinthorne N 2000 Penalized weighted least-squares image reconstruction for dual energy X-ray transmission tomography *IEEE Trans. Med. Imaging* **19** 1075–81
- Thibault J B, Sauer K D, Bouman C A and Hsieh J 2007 A three-dimensional statistical approach to improved image quality for multislice helical CT *Med. Phys.* **34** 4526–44
- Wang J, Li T, Lu H and Liang Z 2006 Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose X-ray computed tomography *IEEE Trans. Med. Imaging* **25** 1272–83
- Wang J, Lu H, Eremina D, Zhang G, Zhang G, Wang S, Chen J, Manzione J and Liang Z 2008 An experimental study on the noise properties of X-ray CT sinogram data in the Radon space *Proc. SPIE Med. Imaging* **6913** 69131M

- Wen N, Guan H, Hammoud R, Pradhan D, Nurusev T, Li S and Movsas B 2007 Dose delivered from Varian's CBCT to patients receiving IMRT for prostate cancer *Phys. Med. Biol.* **52** 2267–76
- Whiting B, Massoumzadeh P, Earl O, O'Sullivan J, Snyder D and Williamson J 2006 Properties of preprocessed sinogram data in x-ray computed tomography *Med. Phys.* **33** 3290–303
- Xing L, Thorndyke B, Schreibmann E, Yang Y, Li T F, Kim G Y, Luxton G and Koong A 2006 Overview of image-guided radiation therapy *Med. Dosim.* **31** 91–112
- Yang Y, Schreibmann E, Li T, Wang C and Xing L 2007 Evaluation of on-board kV cone beam CT (CBCT) based dose calculation *Phys. Med. Biol.* **52** 685–705
- Zhong J, Ning R and Conover D 2004 Image denoising based on multiscale singularity detection for cone beam CT breast imaging *IEEE Trans. Med. Imaging* **23** 696–703
- Zhuang T, Nett B, Leng S and Chen G 2006 A shift-invariant filtered backprojection (FBP) cone-beam reconstruction algorithm for the source trajectory of two concentric circles using an equal weighting scheme *Phys. Med. Biol.* **51** 3189–210
- Zou Y and Pan X 2004 Exact image reconstruction on PI-lines from minimum data in helical cone-beam CT *Phys. Med. Biol.* **49** 941–59
- Zou Y, Pan X and Sidky E 2005 Theory and algorithms for image reconstruction on chords and within regions of interest *J. Opt. Soc. Am. A* **22** 2372–84

A fiducial detection algorithm for real-time image guided IMRT based on simultaneous MV and kV imaging

Weihua Mao, Nadeem Riaz, Louis Lee, Rodney Wiersma, and Lei Xing^{a)}

Department of Radiation Oncology, Stanford University School of Medicine, Stanford, California 94305-5847

(Received 9 January 2008; revised 2 May 2008; accepted for publication 11 June 2008; published 11 July 2008)

The advantage of highly conformal dose techniques such as 3DCRT and IMRT is limited by intrafraction organ motion. A new approach to gain near real-time 3D positions of internally implanted fiducial markers is to analyze simultaneous onboard kV beam and treatment MV beam images (from fluoroscopic or electronic portal image devices). Before we can use this real-time image guidance for clinical 3DCRT and IMRT treatments, four outstanding issues need to be addressed. (1) How will fiducial motion blur the image and hinder tracking fiducials? kV and MV images are acquired while the tumor is moving at various speeds. We find that a fiducial can be successfully detected at a maximum linear speed of 1.6 cm/s. (2) How does MV beam scattering affect kV imaging? We investigate this by varying MV field size and kV source to imager distance, and find that common treatment MV beams do not hinder fiducial detection in simultaneous kV images. (3) How can one detect fiducials on images from 3DCRT and IMRT treatment beams when the MV fields are modified by a multileaf collimator (MLC)? The presented analysis is capable of segmenting a MV field from the blocking MLC and detecting visible fiducials. This enables the calculation of nearly real-time 3D positions of markers during a real treatment. (4) Is the analysis fast enough to track fiducials in nearly real time? Multiple methods are adopted to predict marker positions and reduce search regions. The average detection time per frame for three markers in a 1024×768 image was reduced to 0.1 s or less. Solving these four issues paves the way to tracking moving fiducial markers throughout a 3DCRT or IMRT treatment. Altogether, these four studies demonstrate that our algorithm can track fiducials in real time, on degraded kV images (MV scatter), in rapidly moving tumors (fiducial blurring), and even provide useful information in the case when some fiducials are blocked from view by the MLC. This technique can provide a gating signal or be used for intra-fractional tumor tracking on a Linac equipped with a kV imaging system. Any motion exceeding a preset threshold can warn the therapist to suspend a treatment session and reposition the patient. © 2008 American Association of Physicists in Medicine. [DOI: [10.1118/1.2953563](https://doi.org/10.1118/1.2953563)]

I. INTRODUCTION

Highly conformal radiation therapy techniques, such as three-dimensional conformal radiotherapy (3DCRT) and intensity-modulated radiation therapy¹ (IMRT), provide exquisitely shaped radiation doses that closely conform to tumor dimensions while sparing sensitive structures.^{1,2} They require greater precision in tumor localization, treatment setup, and delivery than conventional techniques. In practice, inter- and intrafraction organ motion results in an uncertainty of tumor location. For example, respiratory and prostate tumors can move up to 3 cm over the course of routine radiotherapy.^{3–9} Research activities on image-guided radiation therapy have emerged recently to improve targeting in radiation treatment.¹⁰ It is essential to track the dynamical nature of human anatomy or at least the tumor motion in real time.¹¹

Several methods of obtaining the real-time tumor position are available, and these can be categorized as being either indirect (external surrogate based) or direct (fiducial/image) in nature. In general, indirect tumor location methods, such as external skin marker tracking or breath monitoring techniques, rely on the correlation between external body param-

eters and the tumor.^{5,12} In reality, the relationship between external parameters and internal organ motion is complex and a large uncertainty may be present in predicting the tumor location based on external markers. Direct tumor position measurement is highly desirable for therapeutic guidance. In the last decade, a number of direct real-time 3D tumor tracking methods have been implemented, primarily using fluoroscopy^{5,11,13} or magnetic field localization.¹⁴ Particularly, the feasibility of using electronic portal imaging devices (EPID) and stereoscopic x-ray imaging for real-time tumor tracking has been explored.^{3,5,6,15–26}

A crucial component of an image based tracking system is the ability to successfully identify and track user-specified image-based features at a near real-time speed. The detection algorithm must also be able to segment markers from anatomic structures and simultaneously track multiple markers without confusing one marker for another. Generally, an intensity based fiducial marker detection algorithm tends to fail when the marker is in the vicinity of high contrast structures such as bone. This can be avoided by using template matching, as demonstrated by Shirato *et al.*, for the tracking of a single spherical gold marker using multiple kV fluoroscopic

imaging systems.²⁵ Tang *et al.* have further extended template matching by developing a cylindrical marker detection algorithm that takes into account the different possible projections of the marker based on its orientation and length on their custom designed stereoscopic kV on-board imaging systems.²⁷ Because a single in-line x-ray beam is only two dimensional, the 3D coordinates of the embedded fiducials are usually obtained by specially designed multiple kV x-ray sets.^{3,5,6,25–28}

Recently, Wiersma *et al.* used combined kV and MV imaging systems to track the 3D location of a spherical metallic fiducial.²⁹ This technique has the inherent benefit in that only one kV source is required for full 3D marker positional information since the actual MV beam is also used for positioning. Compared to other stereoscopic systems, which generally require two or more kV imaging sets, this technique reduces the radiation dose to the patient and requires minimal modification of the current hardware. In Wiersma's work, a freely available third party software program was used to detect a spherical fiducial, 3 mm in diameter. As of yet, there have been few works presenting marker-tracking algorithms that are suitable for tracking internal markers using MV image data with high success rates because these images have significantly reduced contrast.¹⁸ Further, it is a more challenging task to robustly detect small cylindrical fiducials (gold seeds) used clinically, in a realistic setting. This problem is exacerbated when the incident beam is an IMRT field instead of an open field, because one or all implanted fiducials may be blocked by the MLC at some segments during IMRT delivery.

In this article, using a new detection algorithm,³⁰ we will study four clinically relevant issues pertaining to the application of tracking fiducials in real time based on simultaneous kV and MV imaging:

- (1) How fast can a marker move and still be detectable? Motion will blur the fiducial and hinder tracking it. It is essential to investigate the maximum moving speed at which the markers can still be detected by this procedure.
- (2) How does the MV beam scattering affect kV imaging? It has been reported that if MV and kV images are acquired simultaneously, MV beam scattering has significant interference with kV imaging while kV beam scattering effect on MV images is minor.³¹ A quantitative study is necessary to clarify if simultaneous MV beam scattering affects the detection of markers on kV images.
- (3) Can an algorithm reliably track markers in MV images with irregular fields? The aperture of MV beam is carefully modified by a multiple-leaf collimator (MLC) in any 3DCRT or IMRT treatment plans. A major challenge is that one or more markers may be outside of the MV fields/ images and this requires very high specificity detection, particularly when the MLC is moving. Is this tracking procedure capable of handling the MLC blocking problem?
- (4) Is this algorithm fast enough to track markers in nearly

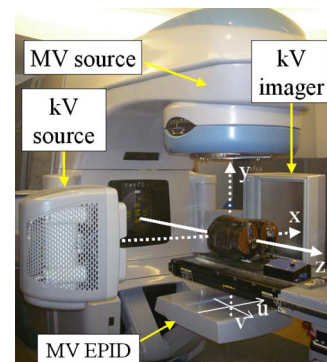


FIG. 1. A picture of the Trilogy with extended MV EPID, kV source, and kV imager. Room coordinates are illustrated.

real time? An analysis speed of 10 Hz is desirable for a nearly real-time 3D position tracking in order to gate treatment beam delivery.

II. MATERIALS AND METHOD

II.A. Hardware setup and computer calculation

All experiments were performed on a Varian TrilogyTM (Varian Medical System, Palo Alto, CA) with a MV EPID and a kV onboard imaging system as shown in Fig. 1. The onboard imaging system is located perpendicular to the treatment MV beam and consists of a kV x-ray tube together with an α Si flat panel imager. Effective pixel sizes of the kV and MV detectors were 0.388 and 0.392 mm, respectively. Both detectors had a resolution of 1024×768 , corresponding to an effective area of detection of approximately $40 \text{ cm} \times 30 \text{ cm}$. For both MV and kV systems, the default source-to-axis distances (SADs) and source-to-imager distances (SIDs) were set to 100 and 150 cm, respectively. Both dual MV energies, 6 and 15 MV, were used in this article. The MV EPID was capable of capturing images at a speed of 7.5 frames per second (fps) and 7.8 fps for 6 and 15 MV beams, respectively, while the kV imager had a capturing speed of 15 fps in the fluoroscopic or continuous imaging mode.

A sliced pelvic phantom was tested on a motion platform. Three gold cylindrical fiducial markers were embedded in the prostate position, each fiducial had a diameter of 1.2 mm and a length of 3.0 mm. The platform was driven by an electrical motor and could move linearly with an adjustable period between 2.0 and 6.0 s and its maximum motion amplitude was set to 1.0 cm.

A software program (C language) was developed to analyze projection images and obtain fiducial positions. All calculations were performed on a Dell Precision 470 workstation (3.4 GHz Xeon CPU and 4 GB RAM).

II.B. Fiducial tracking algorithm

Figure 2 illustrates the complete procedure to track markers on open-field kV or MV images. Major steps are described in the following subsections.

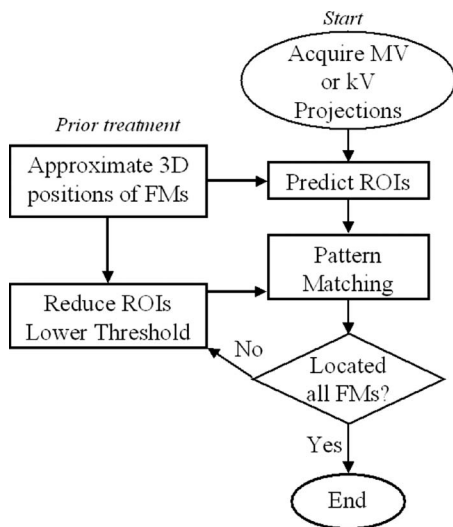


FIG. 2. Flow chart of detecting moving markers on open-field kV/MV images.

- Obtain approximate 3D positions of FMs from planning CT

The first step was to obtain approximate 3D positions of fiducials from the planning CT and then convert them to the treatment coordinates. A simple intensity-based search for the markers was performed on the planning CT. Due to the much larger CT numbers of the metallic markers relative to other anatomical structures, the markers were easily segmented from the image background. The displacement vector relating the CT iso-center to the treatment iso-center was then used to relocate marker CT coordinates relative to the treatment iso-center.

- Predict FM locations

The FM locations on projection images could be predicted by their approximate 3D positions obtained from prior planning CT images. For any FM having 3D position (x_M, y_M, z_M) , its expected projection location (u, v) on either the kV or MV detector was predicted by the following relationships:³²

$$u = F \frac{\cos(\phi)x_M + \sin(\phi)y_M}{R - \sin(\phi)x_M + \cos(\phi)y_M} \quad (1)$$

$$v = F \frac{z_M}{R - \sin(\phi)x_M + \cos(\phi)y_M}, \quad (2)$$

where ϕ was the Varian gantry angle, R was SAD, and F was SID. The x axis was in the lateral direction of patient couch, the y axis was in the anterior-posterior direction, the z axis was the superior-inferior (SI) direction, and the origin was the Linac's iso-center. The coordinates of the imagers (u, v) were defined on the imager and rotated with the gantry while v axis was parallel to z axis and u -axis laying the xy plane. Their origin is at the imager geometric center and their units are in mm or pixels (converted by effective pixel size).

The region to search for the fiducial marker was then

reduced to a small circular region of interest (ROI) centered on the predicted location. The ROI may be large or small depending on the range of tumor motion. Typically, a ROI with a radius of 75 pixels, or about 2 cm around the object, was found to be adequate for locating the moving markers in our studies.

For those markers located near each other, overlapping ROIs were combined into one larger ROI group containing both markers. In the case of multiple markers existing in the same ROI group, the detected marker positions in this ROI group were reordered based on the information from the planning CT results including the internal distances between markers and the shortest distances between detected and predicted marker projections, so that every detected fiducial would be identified without any confusion.

- Reduce ROIs

Two methods were used to reduce ROIs if possible. One method was to use the detected FM locations on the previous image during a continuous imaging course. Due to the short time interval between acquisitions (<0.1 s) marker movement between consecutive images was limited so the new FM location should be close to the prior location. For instance, when the ROI radius was reduced to 25 pixels, this still covered a region with a radius of 6.7 mm, which corresponded to any movement with a speed less than 5 cm/s.

Another method was to use the relative positions among markers to locate undetected markers under the rigid body assumption. As long as one or more markers were detected, other markers' expected positions were calculated based on their relative locations to the detected markers by ignoring the internal motion among markers. The second search was usually performed in smaller ROIs with lower thresholds in order to detect markers with a weaker signal.

- Match patterns

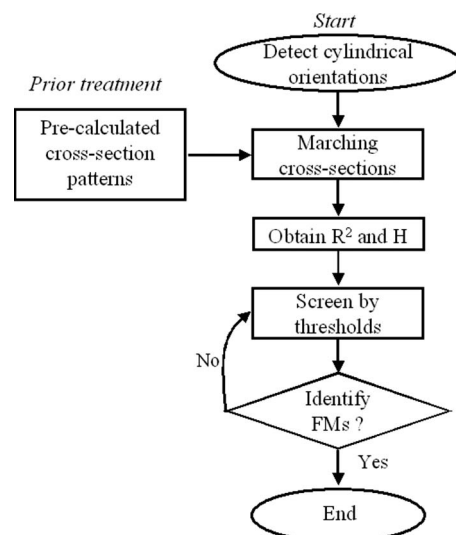


FIG. 3. Flow chart of pattern matching algorithm.

A novel fiducial tracking algorithm (pattern matching algorithm, as summarized in Fig. 3) has been developed to detect cylindrical fiducial markers on kV and MV projection images.^{20,30} A cylindrical marker could be projected into different shapes depending on its orientation. A filter was applied to find the fiducial orientation at a given pixel. With this filter, we divided the angular space on a projection plane into eight bins. The average intensities of every bin adjacent to the given pixel were compared and the bin with the highest average intensity was presumed to correspond to the fiducial orientation. It should be noted that the cylindrical orientation results carried no information for most of the pixels and it was only valid for pixels on a cylindrical object.

A cylindrical marker can be projected into different

$$R_{u,v}^2 = \frac{\sum_{(i,j) \in \text{Pattern}} (f_{u+i,v+j} - \overline{f_{u,v}})(p_{i,j} - \bar{p})}{\sqrt{\sum_{(i,j) \in \text{Pattern}} (f_{u+i,v+j} - \overline{f_{u,v}})^2} \cdot \sqrt{\sum_{(i,j) \in \text{Pattern}} (p_{i,j} - \bar{p})^2}}, \quad (3)$$

$$H_{u,v} = \frac{\sum_{(i,j) \in \text{Pattern}} (f_{u+i,v+j} - \overline{f_{u,v}})(p_{i,j} - \bar{p})}{\sum_{(i,j) \in \text{Pattern}} (p_{i,j} - \bar{p})^2}, \quad (4)$$

where $\overline{f_{u,v}}$ is the average intensity of the pattern region around pixel (u,v) and \bar{p} is the average intensity of the pattern distribution as given by

$$\overline{f_{u,v}} = \frac{1}{N} \sum_{(i,j) \in \text{Pattern}} f_{u+i,v+j}, \quad (5)$$

$$\bar{p} = \frac{1}{N} \sum_{(i,j) \in \text{Pattern}} p_{i,j}, \quad (6)$$

with N being the total pixel number of the pattern. The square of the coefficient of correlation or coefficient of determination for a linear regression is conveniently called R square or R^2 . It varied from 0 (no correlation) to 1 (perfect correlation). The scaling factor, H , indicated the relative intensity of the object compared to the background. In an ideal case, the image was scaled from the pattern after a background shift, $f_{u+i,v+j} = k \cdot p_{i,j} + b$, where k and b were constants, and the pattern matching result would be $R_{u,v}^2 = 1$ and $H_{u,v} = k$.

After the R^2 and H values were calculated for every pixel in the ROIs, they were screened based on predefined thresholds. For pixels that passed the threshold, adjacent pixels were examined, and an overall length for the fiducial was determined. The presumed fiducial

shapes (length and orientation) depending on its orientation and the x-ray beam direction. However, its cross section, a narrow section perpendicular to the cylindrical orientation, is a unique feature and it only depends on the fiducial width and cylindrical orientation. For convenience, the cross-section patterns at all eight cylindrical orientations were precalculated for given widths. The cross section of every pixel was then matched with the predetermined cross-section pattern at the cylindrical orientation of that center pixel. At each pixel location (u,v) within ROIs, a comparison was made between the surrounding pixels ($\{f_{u+i,v+j}\}$) and the pattern ($\{p_{i,j}\}$) corresponding to the cylindrical orientation of that pixel. Two criteria were used to quantify the matching: The square of the correlation coefficient, $R_{u,v}^2$, and the scaling factor or intensity, $H_{u,v}$,

would be rejected if its overall length was longer than a predefined maximum length.

• Calculate 3D coordinates from orthogonal dual projections

The nearly real-time 3D positions of fiducial markers were calculated from FM projections on two orthogonal imagers. Every marker (x_M, y_M, z_M) had two projections: (u_{MV}, v_{MV}) on MV imager and (u_{kV}, v_{kV}) on kV imager. The Trilogy might have a different SID and SAD, (F_{MV}, R_{MV}) and (F_{kV}, R_{kV}) , for the MV and kV imaging systems. Because the Trilogy kV imaging system always had a gantry angle of 90° larger than the MV imaging system, it is convenient to use the MV gantry angle to represent the Trilogy's rotation status. At a MV gantry angle of ϕ , the fiducial marker's coordinates could be calculated from Eqs. (1) and (2):

$$\alpha = u_{MV} \frac{F_{kV} R_{MV} + R_{kV} u_{kV}}{F_{MV} F_{kV} + u_{MV} u_{kV}}, \quad (7)$$

$$\beta = u_{kV} \frac{F_{MV} R_{kV} - R_{MV} u_{MV}}{F_{MV} F_{kV} + u_{MV} u_{kV}}, \quad (8)$$

$$x_{FM} = \cos(\phi) \cdot \alpha - \sin(\phi) \cdot \beta, \quad (9)$$

$$y_{FM} = \sin(\phi) \cdot \alpha + \cos(\phi) \cdot \beta, \quad (10)$$

$$z_{FM} = \frac{1}{2} \left(\frac{R_{MV} + \beta}{F_{MV}} v_{MV} + \frac{R_{kV} - \alpha}{F_{kV}} v_{kV} \right). \quad (11)$$

II.C. Experiments

• Speed of moving FMs

The motion of fiducials would blur images and impede tracking. The maximum detectable moving speed is an important factor of this detection procedure. We tracked fiducial markers implanted in a phantom, which moved at various speeds, on open-field kV and MV images. The maximum linear speed can be calculated by Eq. (12)

$$\text{speed} = \frac{2\pi A}{T}, \quad (12)$$

where A was the motion amplitude and it was 10.0 mm in this study and T was the motion period.

Detection efficiency has been defined as the ratio between total number of detected fiducials and total number of fiducials in all images.

• Test the effects of MV scattering on kV imaging

During simultaneous MV and kV imaging, scattering of the MV beam would significantly diminish kV imaging quality, but the diminished kV image might still be good enough to track fiducial markers. In general, the quality of kV images was inversely proportional to MV beam size and directly proportional to kV SID. In this study, the kV images were analyzed by varying the MV field size from the maximum of $26 \times 20 \text{ cm}^2$ (open field) to $10 \times 10 \text{ cm}^2$ and increasing the SID from 150.0 to 181.8 cm. Since the typical human abdomen has different thickness in lateral (LAT) and anterior-posterior directions and MV and kV beams are in orthogonal directions, the scattering effects on both AP and LAT directions were tested here.

• Tracking fiducials with the presence of a MLC

In order to track fiducials during a treatment course, two phantom verification plans for 3DCRT and IMRT treatments from real patients were prepared on an Eclipse (Varian Medical System, Palo Alto, CA). The 3DCRT plan contains four fields: AP, PA, and two LAT fields. The step and shoot IMRT treatment plan included seven fields and 74 segments in total. All MV and kV images were acquired simultaneously during the delivery of the treatment plans. The pelvic phantom was placed on the motion platform and both plans were delivered multiple times at different motion cycle periods.

Figure 4 shows the flow chart to track markers when the MLC was moving (i.e., for an IMRT plan). An identical tracking method was applied for kV images as described in Sec. II B. The kV results were then used to help locate fiducials on MV images as shown in Fig. 4. Compared to tracking markers on open-field images, there was an additional step to screen and define the MLC field after MV images were acquired. An intensity threshold was computed from the minimum and maximum intensity of the whole image and then used for segmenting the MV field from the MLC blocked area by a simple intensity based screening process. The image was also linearly normalized for consistent analysis

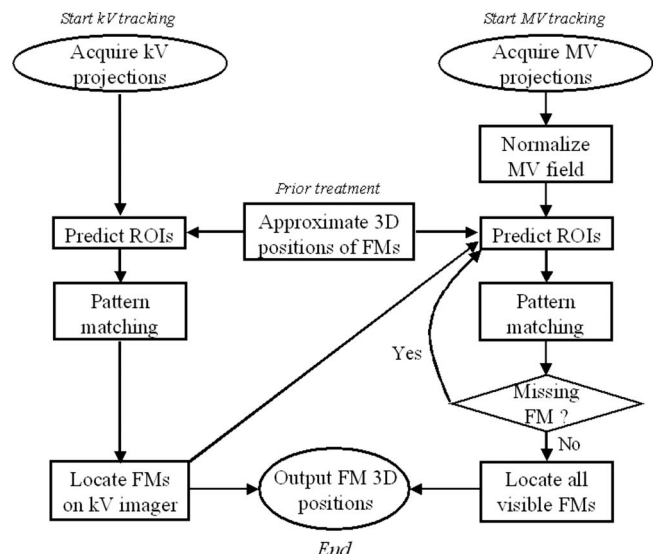


FIG. 4. Flow chart of detecting moving markers on IMRT kV/MV images.

across a series of projections. For convenience, the maximum intensity was normalized to 1000 while the minimum intensity was normalized to 0.

When using the treatment MV beam for in-line imaging, a potential difficulty is that the fiducials may be partially or completely blocked by the MLC at certain segments. There are five sources of information can help in this situation. First, the coordinates of any fiducial from the orthogonal kV imaging are always available. This piece of information is very valuable in estimating the position of this fiducial on MV imager because the kV projection result indicates that the fiducial locates on a kV x-ray trace from the kV source to the kV projection. Its possible projection on MV imager must be on the projection of this kV x-ray trace. Coordinates of kV and MV projections of the same fiducial on the common axis, the v axis, are close to each other. More details are presented in the Appendix. Second, fiducial kinetics obtained at the previous time point can be utilized by a prediction algorithm to facilitate estimation of marker position as discussed in the subsection II B. **Fiducial tracking algorithm, Reduce ROIs.** Third, the detected fiducials' positions could be helpful to locate other blocked fiducials based on their relative positions as discussed in the subsection II B. **Fiducial tracking algorithm, Reduce ROIs.** Fourth, the MLC leaf positions are always available from the EPID images, which can serve as a useful landmark for fiducial position estimation because the fiducials should be either detected or in the MLC leaf blocked area. Finally, the 3D movement of the markers captured by the pre-treatment 4D CBCT and simulation 4D CT is also available for positional prediction. For example, an elliptic ROI may be used for a known tumor- motion direction instead of a circular ROI. This will reduce the size of possible MV projection locations in certain directions. Furthermore, it would reduce the length of 1D possible

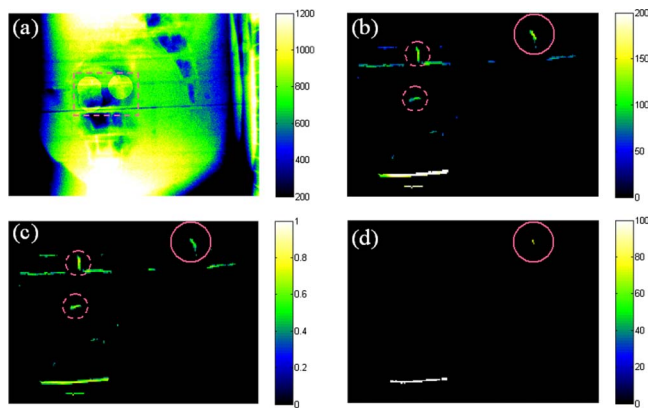


FIG. 5. Typical fiducial tracking. (a) Projection image in which the ROIs were highlighted. (b) Intensity results of pattern matching in the dash rectangle region of (a). (c) R^2 results of pattern matching in the dash rectangle region of (a). (d) Qualified features after the primary search.

MV projection locations predicted by orthogonal kV imaging results.

III. RESULTS

• An Example of fiducial tracking

Figure 5 illustrates a difficult fiducial tracking case on a lateral MV image. The ROIs for three fiducials were predicted by the planning CT results and highlighted in Fig. 5(a), in which the dash rectangle displayed a region including all ROIs. After pattern matching was performed on every pixel in the ROIs, intensity and R^2 results were obtained as shown in Figs. 5(b) and 5(c), respectively. The intensity and R^2 thresholds were determined based on the maximum values of that ROI group and predefined minimum values. After threshold screening, only one feature was qualified as a fiducial [enclosed by a solid circle in Fig. 5(d)] while another feature was rejected because its length was too long.

The second search then followed. The positions of undetected markers were closely estimated by the relative shifts among markers as enclosed by the small dashed circles in Figs. 5(b) and 5(c). After the thresholds were set at lower values, all markers were successfully detected.

• Speed limit for detecting moving fiducials

MV and kV image series were analyzed with the motion platform moving at different periods. Figure 6(a) illustrates the projection locations of one marker when the motion platform was moving at an amplitude of 10.0 mm with a period of 4 s. The traces were fit to sine waves shown in the Fig. 6(a). The fitting results indicated that the motion amplitudes for both axes were 3.38 and 9.49 mm, respectively. This means the total amplitude is 10.07 mm ($=\sqrt{3.38^2+9.49^2}$), which is very close to the nominal amplitude (10.0 mm). Figure 6(b) illustrates the difference between measured projections and fitting results and the variations are less than 0.8 mm.

Images of the moving phantom were analyzed to test

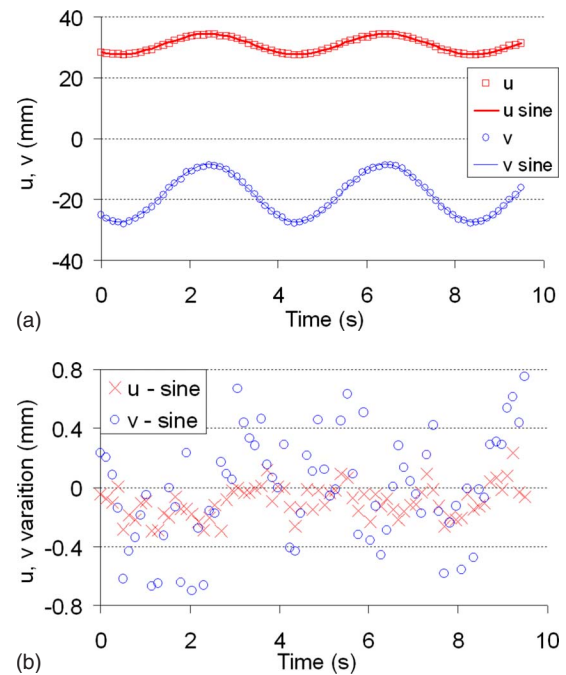


FIG. 6. Marker Detecting Accuracy. (a) Measured the projections against fitting results. (b) Differences between measured the projections against fitting results.

the detection efficiency for moving fiducials. In total, 877 kV and 456 MV images on both lateral (90° gantry angle) and AP directions were acquired and analyzed for the motion study. It is always easier to detect markers on the AP images than on the LAT images, partially because of the increased scattering at the LAT direction. All markers in kV images and AP MV images (either 6 MV or 15 MV energy) were successfully detected (detection efficiency of 100%). For the difficult cases, detection efficiencies on LAT MV images are illustrated in Fig. 7. A detection efficiency of better than 95% is achieved for both MV energies when the motion period is 4 s or longer, which correlates to a maximum linear speed of 1.6 cm/s or slower.

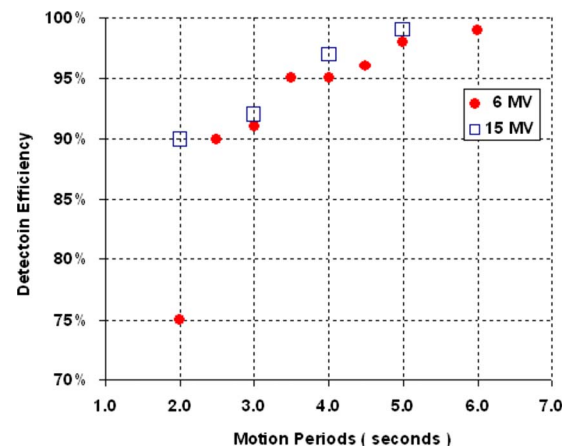


FIG. 7. Detection efficiencies of moving markers on LAT MV images of dual energies.

TABLE I. Summary of the kV detection efficiencies affected by MV field size and kV SID. Two efficiencies present MV field at AP and LAT directions, respectively.

kV SID (cm)		181.8		166.7		150.0	
MV direction		AP	LAT	AP	LAT	AP	LAT
MV field size (cm ²)	10×10	100%	100%	100%	100%	100%	100%
	15×15	100%	100%	100%	100%	100%	4%
	26×20	100%	76%	61%	13%	6%	1%

• Effects of MV beam scattering on kV imaging

Usually scattering of the simultaneous MV and kV beams can cause concern for degrading the quality of the kV image. Fortunately, here only the fiducial markers are of interest and our analysis is capable of suffering a certain degree of scattering. The kV SID was extended further to reduce the scattering from MV beams; 842 kV images were acquired and analyzed while varying the SID and MV field size; the detection efficiency results are summarized in Table I. As expected, either a smaller MV field or a larger SID leads to better detection. It was found that the markers in kV images could be successfully analyzed even with the scattering from a common treatment MV field (10×10 cm or less). In addition, when our 3DCRT and IMRT plans were delivered, more than 5000 kV images were acquired and analyzed and all markers were successfully detected with a detection efficiency of 100%. This guarantees that kV imaging is always applicable for tracking fiducials during treatment.

• Tracking FMs in the presence of a MLC

The IMRT plan was delivered once without phantom motion and once with a motion period of 4.0 s. All MV and kV images were acquired and analyzed. Special attention was paid on treatment segments with fiducials

partially or completely blocked by the MLC. Figure 8 illustrates fiducial tracking on the MV field at a gantry angle of 235° with the static phantom. The fiducial coordinates obtained from orthogonal kV images were used to help locate projections on MV images [as shown in Fig. 8(a)]. The possible MV projection might not be exactly horizontal due to the divergence of the kV x-ray trace (as discussed in the Appendix). As long as a fiducial was detected at any segment [Figs. 8(c)–8(i)], the blocked fiducials could be estimated very closely by fiducial relative positions (circles in the MLC blocked area) and common axis kV coordinates (solid lines). The estimated position was limited on a horizontal line with a length corresponding internal deformation up to ± 5 mm. It is still possible that all fiducials are outside of the MV field because of small IMRT segments as shown in Fig. 8(b). In this case, the possible positions could be estimated as elliptic ROIs by 3D planning CT data and available 4D CT results. Their long and short axes correspond to possible motion of ± 2 cm and ± 1 cm in two directions, respectively. The orthogonal kV imaging results limit the estimation to horizontal lines with a length corresponding motion up to ± 1 cm.

The 3DCRT plan was delivered three times with the motion period varied at 2.0, 3.0, and 4.0 s. All MV and kV images were acquired and analyzed. Figure 9 displays the simultaneous MV and kV images when the phantom was moving. Most of the markers were suc-

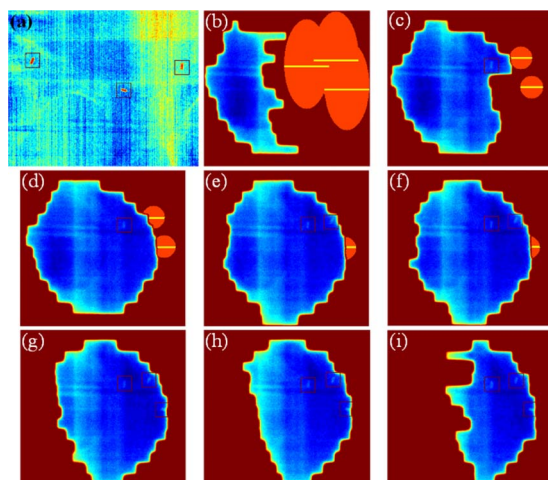


FIG. 8. Detect fiducials during an IMRT treatment. (a) Orthogonal kV image. (b)–(i) MV images for every segment. Detected fiducials were enclosed in square boxes. Ovals or circles in the MLC blocked region indicate the predicted fiducial projections and the horizontal line indicate the predicted fiducial projections based on orthogonal kV imaging results.

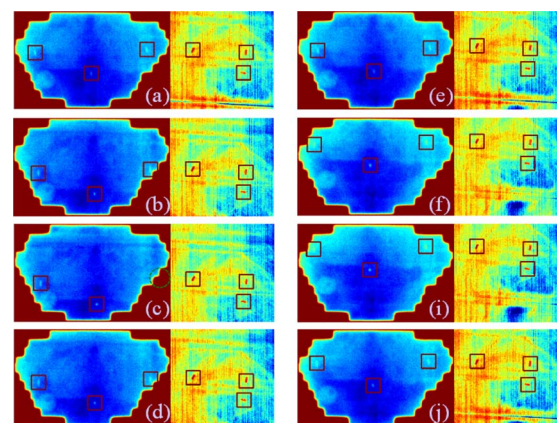


FIG. 9. Simultaneous MV and kV images with a MV gantry angle of 180° and a motion period of 4 s.

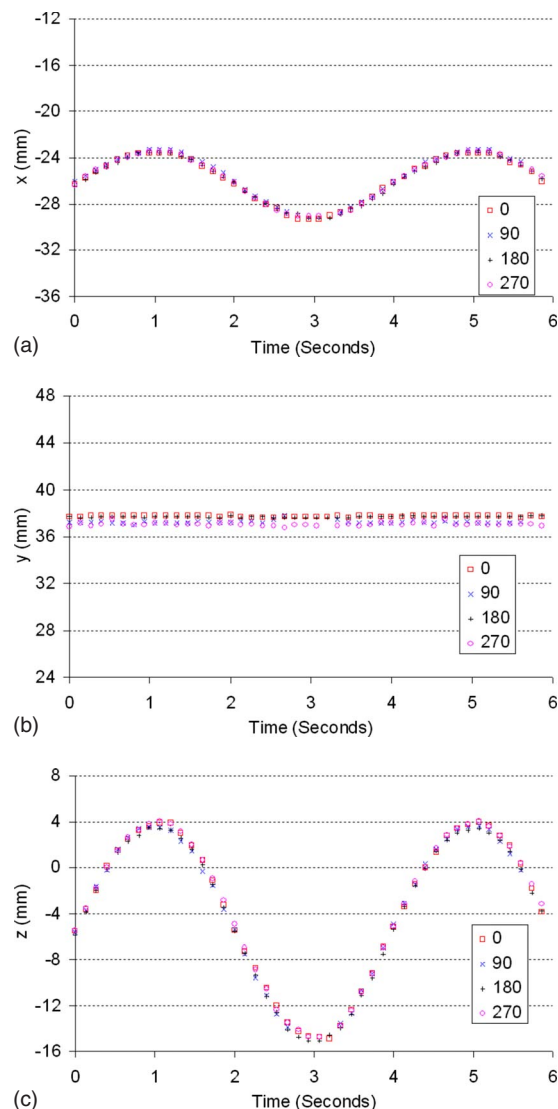


FIG. 10. Three-dimensional coordinates of one marker during the treatment. Results from four fields were synchronized for comparison purpose; (a) x coordinates, (b) y coordinates, (c) z coordinates.

cessfully detected as enclosed in solid rectangles. Nearly real-time 3D positions of fiducial markers were calculated at all fields based on Eqs. (7)–(11). After synchronizing the results from different fields, it was found that results from different fields were within ± 0.8 mm of their average positions. Figure 10 illustrates 3D coordinates of one moving marker in every field with a period of 4 s.

It should be noted that the presence of a MLC does not affect the detection efficiency dependence on motion speed and all fiducial markers on every kV image were successfully detected.

• Performance of analysis software

Special attention was paid on the analysis software programming. It is essential to compute the location of a marker quickly so to be able to track the markers at a speed of 10 fps or higher. Most of the calculation time was spent on the pattern matching, including determin-

ing orientation and calculating intensity and R^2 . A complete search on 1024×768 pixels image costs about 1 s. The searching region of three markers with a radius of 75 pixels reduces the calculation time by an order of magnitude. The average search time for three markers was 0.08 s per frame.

Basically, kV and MV images could be analyzed independently. It is generally much easier to track markers on kV images than on MV images not only because kV images have better contrast but also because fiducials are always visible in kV images. However, this does not mean that double the calculation time is needed for simultaneous kV and MV image analysis because the kV tracking results might be used to reduce MV ROIs as shown in Fig. 4 in addition to the other two methods to reduce ROIs. This process significantly reduces the size of ROIs on MV images and further reduces the computation time. It is found that 0.1 s analysis time per kV/MV pair is achievable on our current single CPU architecture. Reduced computation time is expected if parallel computation is implemented on a multiple-core processor.

IV. DISCUSSION

The four studies presented in this article pave the way to track 3D positions of multiple moving fiducial markers during treatment in nearly real time. We demonstrate that our algorithm is capable of tracking fast moving fiducial markers by simultaneous kV and treatment MV imaging. Furthermore, the analysis speed is fast enough to report nearly real-time 3D positions of the markers. It is also worthwhile to mention that this procedure does not need any assumed motion model. Although all motion experiments were performed on a periodic moving platform, the algorithm does not take this periodicity into consideration, meaning these results should just as easily apply to nonperiodic motion. This is also why the term nearly real-time 3D position was used instead of 4D position because the terminology of 4D usually refers to a periodic motion particularly when it is associated with 4D CT. This procedure detects any regular, irregular, or sudden motion as long as the maximum speed is not very fast.

We can track marker motion up to a period of 2 s in kV images, not only because of much higher contrast of kV images than MV images, but also because of its higher capture speed (15 fps). Currently, the acquisition speed of the MV imaging is less than 8 fps, which means that the marker may move about 2 mm (~ 7.5 pixels on imager) at a speed of 1.6 cm/s during one acquisition (marker width ~ 5 pixels). The blurring induced by motion and small size of the marker make it difficult to track fiducials on MV images, particularly those with low image quality. It should also be noted that better hardware will improve performance, e.g., quicker MV image acquisition.

This method is suitable for tracking any tumor motion as long as fiducial markers or surgical clips are available. It is not limited to a certain type of site although the studies were performed on a pelvic phantom.

The proposed procedure will be an economic solution to nearly real-time tracking of tumor motion since more and more linear accelerators are coming equipped with an on-board kV imaging system. This procedure does not need any extra expensive equipment such as multiple kV imaging sets.

Another advantage of this procedure over stereoscopic kV tracking systems is that using the treatment beam as part of tracking reduces extra dose to the patient. Although kV imaging dose is still a concern, we would like to emphasize that the patient dose from a single kV imaging set here is 50% less than that from stereoscopic kV imaging sets. In this study, we used the typical kV CBCT imaging parameters (125 kV, 80 ms, and 25 mA). It has been reported that the dose to the soft tissue for such a CBCT scan (with 630 images) is ~ 5 cGy.³³ Since an acquisition speed of 7.5 fps is sufficient, the kV dose is about 3.6 cGy/Min ($5 \text{ cGy}/630 \times (7.5/\text{s}) = 0.06 \text{ cGy/s}$). A typical IMRT treatment has a beam on time of about 2 min for a 200 cGy dose delivery to the target. This implies that the kV imaging dose is about 7.2 cGy. However, this represents the worst case scenario where the kV beam is on throughout the IMRT beam delivery process. In reality, it is not necessary to keep the kV imaging system on all the time. The kV imaging may be switched off during the step mode of IMRT delivery. In addition, the kV field size can be reduced according to the fiducial motion range. All these strategies may further reduce the kV imaging dose.

The proposed hybrid MV/kV imaging technique is readily applicable to facilitate conventional 3D radiation therapy by providing real-time information of the implanted fiducials. When using IMRT MV beam for in-line imaging, a potential difficulty is that the fiducials may be partially or completely blocked by the MLC leaves at a certain segment(s). There are four sources of information that help to estimate the 3D coordinates of the MLC-blocked fiducial in this situation. First, the coordinates of the fiducial in the plane perpendicular to the kV beam are still available from kV imaging. This piece of information is very valuable because it significantly reduces the dimensionality of the problem. Second, the fiducial kinetics attained by the kV/MV system at earlier time when the marker(s) are not blocked can be utilized to adaptively predict the “missing” coordinate of the marker in combination with the kV information. Third, the MLC leaf positions are always available from the EPID images, which can serve as a useful landmark for fiducial position estimation. Finally, the 3D movement of the markers captured by the pretreatment CBCT and planning CT is also available as *a priori* knowledge for better positional estimation. The development of such a multiple input single output adaptive prediction algorithm is still in progress. Because there is only one coordinate that needs to be estimated for a short interval of time, we foresee no major difficulty in accomplishing an accurate positional estimation. This remains true in a rare

situation when all the fiducials are blocked by a MLC segment. For certain types of tumors that deform little during the treatment process (e.g., the prostate), the positions of unblocked fiducials can be employed as landmarks in allocating the MLC-blocked fiducial(s).

During our 3DCRT verification plan delivery, at least two of three fiducials were visible in every field even when the phantom was on a motion platform. All fiducials were visible in 95% of the segments. During the delivery of the verification IMRT plan, at least one fiducial was visible in about 75% of the segments. This percentage is highly dependent on the case and the treatment planning system. To increase the visible time of a fiducial, it is possible to add one or more imaging segments (i.e., a segmented field with small monitor unit for the purpose of imaging the fiducials) to the treatment IMRT leaf sequences to facilitate the detection of a fiducial at a certain point of IMRT delivery. Another avenue of research is to take the fiducial information into consideration during the IMRT inverse planning process. With the development of segment-based dose optimization methods,^{34,35} it should be feasible to ensure the visibility of at least one fiducial during the inverse planning process. Of course, the addition of this type of constraint in inverse planning may compromise the achievable dose distribution. But it is arguable that the trade-off will likely be minimal because, after all, the fiducials are all inside the tumor target volume and represent high dose points. We are currently actively studying this issue and the results will be reported elsewhere.

V. CONCLUSION

The four studies presented in this article pave the way to track nearly real-time 3D positions of multiple moving fiducial markers during treatment. It is demonstrated that this proposed process is capable of tracking fast moving fiducial markers (up to 1.6 cm/s) by analyzing simultaneous kV and treatment MV images. The analysis speed is as fast as 10 fps and can report nearly real-time 3D positions of markers with submillimeter accuracy.

ACKNOWLEDGMENTS

This work was supported in part by grants from the Department of Defense (PC040282), the National Cancer Institute (1R01 CA104205), and the Komen Breast Cancer Foundation (BCTR0504071).

APPENDIX

Any fiducial projects on kV imager at P , (u_{kV}, v_{kV}) . This fiducial must be on the line of KP (as shown in Fig. 11):

$$\begin{cases} x = R_{kV} - \gamma \cdot F_{kV} \\ y = \gamma \cdot u_{kV} \\ z = \gamma \cdot v_{kV} \end{cases} \quad (\text{A1})$$

While γ is arbitrary, F_{kV} and R_{kV} are the source-imager distance (SID) and the source-axis distance (SAD) of kV system, respectively.

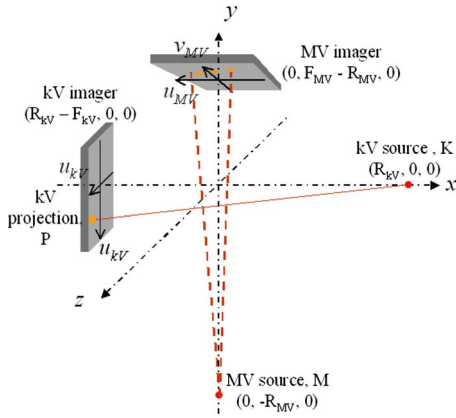


FIG. 11. Schematic diagram of relationship between kV and MV projections of the same fiducial.

This fiducial projects on MV imager on plane MKP (as shown in Fig. 11),

$$x + b \cdot (y + R_{MV}) + c \cdot z = 0, \quad (A2)$$

while b and c are constants and R_{MV} is the SAD of MV system. Input Eq. (A1) to Eq. (A2), we have

$$(R_{kV} - \gamma \cdot F_{kV}) + b \cdot \gamma \cdot u_{kV} + b \cdot R_{MV} + c \cdot \gamma \cdot v_{kV} = 0$$

or

$$R_{kV} + b \cdot R_{MV} + \gamma \cdot (-F_{kV} + b \cdot u_{kV} + c \cdot v_{kV}) = 0. \quad (A3)$$

Since γ is arbitrary, we get

$$b = -\frac{R_{kV}}{R_{MV}} \quad (A4)$$

and

$$-F_{kV} + b \cdot u_{kV} + c \cdot v_{kV} = 0,$$

or

$$c = \frac{F_{kV}R_{MV} + R_{kV}u_{kV}}{R_{MV}v_{kV}}. \quad (A5)$$

The projection on MV imager $[B, (u_{MV}, v_{MV})]$ lies on a line in the MV imager plane $v = F_{MV} - R_{MV}$, so that the projection locations of A and B have a relationship:

$$\frac{v_{MV}}{v_{kV}} = \frac{F_{MV}R_{kV} + R_{MV}u_{MV}}{F_{kV}R_{MV} + R_{kV}u_{kV}}$$

or

$$(F_{kV}R_{MV} + R_{kV}u_{kV}) \cdot v_{MV} - R_{MV}v_{kV}u_{MV} = F_{MV}R_{kV}v_{kV}. \quad (A6)$$

Typically, both MV and kV systems have a SAD of 100 cm and a SID of 150 cm, approximately,

$$\frac{v_{MV}}{v_{kV}} \approx 1 + \frac{u_{kV} - u_{MV}}{F_{MV}}.$$

For a small target, whose fiducials are within a region with a

radius of 2 cm, $|v_{MV}/v_{kV}|$ is usually very close to unity within $\pm 4\%$, i.e., v_{MV} and v_{kV} are close to each other.

^{a)} Author to whom correspondence should be addressed. Electronic mail: lei@reyes.stanford.edu

¹IMRT Collaborative Working Group, "Intensity-modulated radiotherapy: Current status and issues of interest," *Int. J. Radiat. Oncol., Biol., Phys.* **51**(4), 880–914 (2001).

²J. M. Galvin et al., "Implementing IMRT in clinical practice. A joint document of the American Society for Therapeutic Radiology and Oncology and the American Association of Physicists in Medicine," *Int. J. Radiat. Oncol., Biol., Phys.* **58**(5), 1616–1634 (2004).

³K. Kitamura et al., "Three-dimensional intrafractional movement of prostate measured during real-time tumor-tracking radiotherapy in supine and prone treatment positions," *Int. J. Radiat. Oncol., Biol., Phys.* **53**(5), 1117–1123 (2002).

⁴J. M. Crook et al., "Prostate motion during standard radiotherapy as assessed by fiducial markers," *Radiother. Oncol.* **37**(1), 35–42 (1995).

⁵R. I. Berbeco et al., "Residual motion of lung tumors in end-of-inhale respiratory gated radiotherapy based on external surrogates," *Med. Phys.* **33**(11), 4149–4156 (2006).

⁶Y. Seppenwoolde et al., "Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **53**(4), 822–834 (2002).

⁷Y. Xie et al., "Intrafraction motion of the prostate in cyberknife hypofractionated radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* (submitted).

⁸K. L. Zhao et al., "Evaluation of respiratory-induced target motion for esophageal tumors at the gastroesophageal junction," *Radiother. Oncol.* **84**(3), 283–289 (2007).

⁹H. Shirato et al., "Speed and amplitude of lung tumor motion precisely detected in four-dimensional setup and in real-time tumor-tracking radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **64**(4), 1229–1236 (2006).

¹⁰L. Xing et al., "Overview of image-guided radiation therapy," *Med. Dosim.* **31**(2), 91–112 (2006).

¹¹M. J. Murphy, "Fracking moving organs in real time," *Semin. Radiat. Oncol.* **14**(1), 91–100 (2004).

¹²P. C. Chi et al., "Relation of external surface to internal tumor motion studied with cine CT," *Med. Phys.* **33**(9), 3116–3123 (2006).

¹³H. Shirato et al., "Real-time tumour-tracking radiotherapy," *Lancet* **353**(9161), 1331–1332 (1999).

¹⁴P. Kupelian et al., "Multi-institutional clinical experience with the Calypso System in localization and continuous, real-time monitoring of the prostate gland during external radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **67**(4), 1088–1098 (2007).

¹⁵A. L. Boyer et al., "A review of electronic portal imaging devices (EPIDs)," *Med. Phys.* **19**(1), 1–16 (1992).

¹⁶L. Dong et al., "Verification of radiosurgery target point alignment with an electronic portal imaging device (EPID)," *Med. Phys.* **24**(2), 263–267 (1997).

¹⁷M. G. Herman et al., "Technical aspects of daily online positioning of the prostate for three-dimensional conformal radiotherapy using an electronic portal imaging device," *Int. J. Radiat. Oncol., Biol., Phys.* **57**(4), 1131–1140 (2003).

¹⁸P. J. Keall et al., "On the use of EPID-based implanted marker tracking for 4D radiotherapy," *Med. Phys.* **31**(12), 3492–3499 (2004).

¹⁹H. Dehnad et al., "Clinical feasibility study for the use of implanted gold seeds in the prostate as reliable positioning markers during megavoltage irradiation," *Radiother. Oncol.* **67**(3), 295–302 (2003).

²⁰A. Nederveen, J. Lagendijk, and P. Hofman, "Detection of fiducial gold markers for automatic on-line megavoltage position verification using a marker extraction kernel (MEK)," *Int. J. Radiat. Oncol., Biol., Phys.* **47**(5), 1435–1442 (2000).

²¹A. J. Nederveen, J. J. Lagendijk, and P. Hofman, "Feasibility of automatic marker detection with an a-Si flat-panel imager," *Phys. Med. Biol.* **46**(4), 1219–1230 (2001).

²²J. Pouliot et al., "(Non)-migration of radiopaque markers used for on-line localization of the prostate with an electronic portal imaging device," *Int. J. Radiat. Oncol., Biol., Phys.* **56**(3), 862–866 (2003).

²³D. C. Schiffrer et al., "Daily electronic portal imaging of implanted gold seed fiducials in patients undergoing radiotherapy after radical prostatectomy," *Int. J. Radiat. Oncol., Biol., Phys.* **67**(2), 610–619 (2007).

- ²⁴L. Smith *et al.*, "Automatic detection of fiducial markers in fluoroscopy images for on-line calibration," *Med. Phys.* **32**(6), 1521–1523 (2005).
- ²⁵H. Shirato *et al.*, "Physical aspects of a real-time tumor-tracking system for gated radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **48**(4), 1187–1195 (2000).
- ²⁶T. R. Willoughby *et al.*, "Evaluation of an infrared camera and X-ray system using implanted fiducials in patients with lung tumors for gated radiation therapy," *Int. J. Radiat. Oncol., Biol., Phys.* **66**(2), 568–575 (2006).
- ²⁷X. Tang, G. C. Sharp, and S. B. Jiang, "Fluoroscopic tracking of multiple implanted fiducial markers using multiple object tracking," *Phys. Med. Biol.* **52**(14), 4081–4098 (2007).
- ²⁸K. Kitamura *et al.*, "Tumor location, cirrhosis, and surgical history contribute to tumor movement in the liver, as measured during stereotactic irradiation using a real-time tumor-tracking radiotherapy system," *Int. J. Radiat. Oncol., Biol., Phys.* **56**(1), 221–228 (2003).
- ²⁹R. D. Wiersma, W. Mao, and L. Xing, "Combined kV and MV imaging for real-time tracking of implanted fiducial markers," *Med. Phys.* **35**(4), 1191–1198 (2008).
- ³⁰W. Mao, R. Wiersma, and L. Xing, "Fast internal marker tracking algorithm for onboard MV and kV imaging systems," *Med. Phys.* **35**(5), 1942–1949 (2008).
- ³¹W. Luo *et al.*, "Effect of Mv scatter on Kv image quality during simultaneous Kvmv imaging," *Int. J. Radiat. Oncol., Biol., Phys.* **69**(3), S671 (2007).
- ³²W. Mao *et al.*, "CT image registration in sinogram space," *Med. Phys.* **34**(9), 3596–3602 (2007).
- ³³N. Wen *et al.*, "Dose-delivered from Varian's CBCT to patients receiving IMRT for prostate cancer," *Phys. Med. Biol.* **52**(8), 2267–2276 (2007).
- ³⁴D. M. Shepard *et al.*, "Direct aperture optimization: A turnkey solution for step- and-shoot IMRT," *Med. Phys.* **29**(6), 1007–1018 (2002).
- ³⁵C. Cotrutz and L. Xing, "Segment-based dose optimization using a genetic algorithm," *Phys. Med. Biol.* **48**(18), 2987–2998 (2003).

The use of EPID-measured leaf sequence files for IMRT dose reconstruction in adaptive radiation therapy

Louis Lee, Weihua Mao, and Lei Xing^{a)}

Department of Radiation Oncology, Stanford University School of Medicine, Stanford, California 94305

(Received 21 April 2008; revised 23 July 2008; accepted for publication 26 August 2008; published 15 October 2008)

For intensity modulated radiation treatment (IMRT) dose reconstruction, multileaf collimator (MLC) log files have been shown applicable for deriving delivered fluence maps. However, MLC log files are dependent on the accuracy of leaf calibration and only available from one linear accelerator manufacturer. This paper presents a proof of feasibility and principles in (1) using an amorphous silicon electronic portal imaging device (aSi-EPID) to capture the MLC segments during an IMRT delivery and (2) reconstituting a leaf sequence (LS) file based on the leaf end positions calculated from the MLC segments and their associated fractional monitor units. These EPID-measured LS files are then used to derive delivered fluence maps for dose reconstruction. The developed approach was tested on a pelvic phantom treated with a typical prostate IMRT plan. The delivered fluence maps, which were derived from the EPID-measured LS files, showed slight differences in the intensity levels compared with the corresponding planned ones. The dose distribution calculated with the delivered fluence maps showed a discernible difference in the high dose region when compared to that calculated with the planned fluence maps. The maximum dose in the former distribution was also 2.5% less than that in the latter one. The EPID-measured LS file can serve the same purpose as a MLC log files does for the derivation of the delivered fluence map and yet is independent of the leaf calibration. The approach also allows users who do not have access to MLC log files to probe the actual IMRT delivery and translate the information gained for dose reconstruction in adaptive radiation therapy. © 2008 American Association of Physicists in Medicine. [DOI: [10.1118/1.2990782](https://doi.org/10.1118/1.2990782)]

Key words: EPID, leaf sequence file, IMRT, dose reconstruction

I. INTRODUCTION

An advantage of the fractionation scheme in radiation treatment is that it offers room for adaptive radiation therapy (ART). ART is a radiation treatment strategy of which the subsequent fractional delivery can be adaptively modified based on a closed-loop control framework using systematic feedback of geometric and dosimetric information.¹⁻³ The ultimate goal of ART is to maintain adequate target coverage with a desired dose and ensure doses received by normal tissue are within tolerance at the conclusion of treatment. The adaptive strategy comes into play at different levels of complexity depending on the techniques and resources available. It ranges from the most accessible form of adapting treatment margins based on daily portal images⁴ to the most sophisticated one of reoptimization or replanning of treatment plans.^{1,2} Common to all these strategies is the execution of dose reconstruction at some stage during the ART process. Through this, the dose deposited to a patient in a particular fraction can be correlated or mapped to a reference set of computed tomography, ideally by deformable registration, contributing to an accumulated dose delivered so far, which is a key parameter for the feedback mechanism in the ART framework.

However, in most intensity modulated radiation treatments (IMRTs) employing ART strategies, the dose reconstruction is tacitly based on an assumption that the delivery

of fluence maps is as planned.⁵⁻⁷ This assumption might not be necessarily valid. For instance, in IMRT using step-and-shoot mode, the expected delivery of fluence maps might not be realized due to intrinsic errors associated with the multileaf collimator (MLC) kinematics and beam control communication resulting in overshoot, undershoot segmental monitor units, dropped segments, and beam delivery during leaf motion.⁸⁻¹⁰ In order to incorporate these errors in the dose reconstruction, Lee *et al.*¹¹ and Litzenberg *et al.*¹⁰ have demonstrated a pragmatic approach of using MLC log files to reconstruct the IMRT dose actually delivered. This is based on the fact that the MLC log files have been validated to faithfully reflect the actual delivery process of MLC-based IMRT.^{12,13} Because the MLC log file is only available from one commercial linear accelerator (linac) manufacturer (Varian Medical Systems, Palo Alto, CA), users with linacs from other manufacturers are deprived of this straightforward approach to reconstruct the delivered IMRT dose. Furthermore, the leaf position data recorded in a MLC log file is taken from the same encoders used to position the leaves, making the reported position dependent on the leaf position calibration and by no means an absolute measure of the leaf position. Any systematic error introduced in the MLC calibration might lead to actual leaf positions different from the expected ones, resulting in dose errors.¹⁴

In order to circumvent this dependence and provide a universal approach of probing the actual delivery of a fluence

map, we propose using an amorphous silicon electronic portal imaging device (aSi-EPID) to capture every segment of the fluence map during the treatment. For each captured segment, the leaf positions for each pair of leaves are found by an edge detection algorithm; the fractional monitor unit (fMU) associated with this particular segment is also sampled. After all the segments have been analyzed, a leaf sequence (LS) file is reconstituted using the segmental leaf positions and their associated fMU based on the sequence the segments are delivered. The EPID-measured LS files can then be loaded to the treatment planning system (TPS) to derive the delivered fluence maps and reconstruct the delivered IMRT dose. aSi-EPIDs are geometrically and functionally stable, giving undistorted images of high resolution and contrast.^{15,16} The proposed approach is based on the fact that the use of the aSi-EPID in measuring leaf end positions to a high degree of accuracy has been proven, leading to its widespread applications in MLC quality assurance,^{17,18} leaf calibrations,^{19,20} and leaf motion tracking.^{14,21}

The objective of this work is twofold: (1) To present a proof of feasibility and principles in reconstituting an EPID-measured LS file which serves the same purpose for deriving the delivered fluence map as a MLC log file does and (2) to demonstrate the dose reconstruction essential for adaptive radiation therapy using the delivered fluence maps literally calculated from the MLC segments captured by an EPID during an IMRT delivery.

II. METHODS AND MATERIALS

II.A. Description of the MLC and EPID

All experiments were done on a Trilogy linac (Varian Medical Systems, Palo Alto, CA) equipped with a Millennium 120-leaf MLC and kilovoltage/megavoltage EPIDs. The Millennium 120-leaf MLC consists of two banks of 60 leaves. The leaf widths for the central 40 leaf pairs and the outer 10 leaf pairs are 0.5 and 1.0 cm, respectively. The leaves can travel a maximum of 16.5 cm across the beam central axis, and the maximum leaf span between the two leaves on the same carriage is 14.5 cm. All measurements are referred to the isocentric plane. The leaf calibration procedure recommended by Graves *et al.*²² was performed to ensure that the MLC indicated field edge positions agreed with the radiation field edges to within 0.3 mm before the experiments.²³

The megavoltage (MV) EPID (Varian aS1000 flat panel detector) was used to acquire images for the experiments. The EPID is mounted on retractable arms attached to the gantry. It has a detector area of $40 \times 30 \text{ cm}^2$ with a matrix of 1024 by 768 pixels, resulting in a physical pixel size of 0.392 mm. The EPID consists of: (1) A 1.0-mm-thick copper plate for build-up, (2) a phosphor screen of gadolinium oxysulphide doped with terbium (Kodak Lanex Fast Screen) to convert incident radiation to visible photons, (3) a pixel array implanted on an amorphous silicon substratum where each pixel is made up of a photodiode and thin film transistor to convert the light photons to electric charges, and (4) elec-

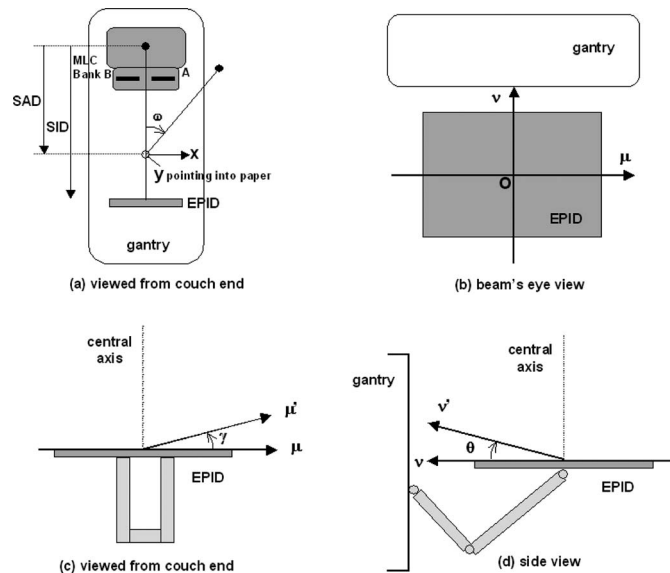


FIG. 1. Coordinate systems describing the imaging geometry of the EPID. (a) Isocentric plane (x, y) and gantry angle ω as viewed from the couch end. (b) EPID (μ, ν) plane as viewed in the beam's eye view. (c) EPID's μ -axis tilt γ as viewed from the couch end. (d) EPID's ν -axis tilt θ as viewed from the side. SAD: Source-to-axis distance; SID: Source-to-imager distance.

tronics for readout. The electrical signals are digitized by a 14 bit analog-to-digital converter and processed into image data.

II.B. Geometric status of the EPID

Baker *et al.*¹⁹ and Parent *et al.*¹⁴ reported that a systematic tilt of the EPIDs was observed in their studies and indicated that it was likely to occur for all different EPIDs; Clarke and Budgell²⁰ have also demonstrated the effect of the gantry angle on the EPID sag. Therefore, it is expected that the imaging geometry for the EPID at different gantry angles might deviate from an ideal configuration that we use as the basis for the measurement of the leaf end position. We need to establish the geometric status of the EPID before we can accurately measure the leaf end position from an EPID image. Recently, our group has developed a geometric quality assurance phantom and an automated analysis program (gQA tool) to study the geometric integrity of the on-board imager.²⁴ This gQA tool was used in this work to study the changes in the overall imaging geometry of the EPID, including the source-to-imager distance (SID), the EPID center, and the tilt of the EPID for every 10° of a full rotation of the gantry. Based on the geometric information provided by the gQA tool, the maximum discrepancies for the SID, EPID center offset in either direction, and tilt were found to be 2.7 mm, 2.2 mm, and 1.4° , respectively. Positional corrections were incorporated into the measurement of the leaf end position.

II.C. Measurement of the leaf end position

The coordinate systems used to describe the imaging geometry of the EPID are shown in Fig. 1. The (x, y) plane

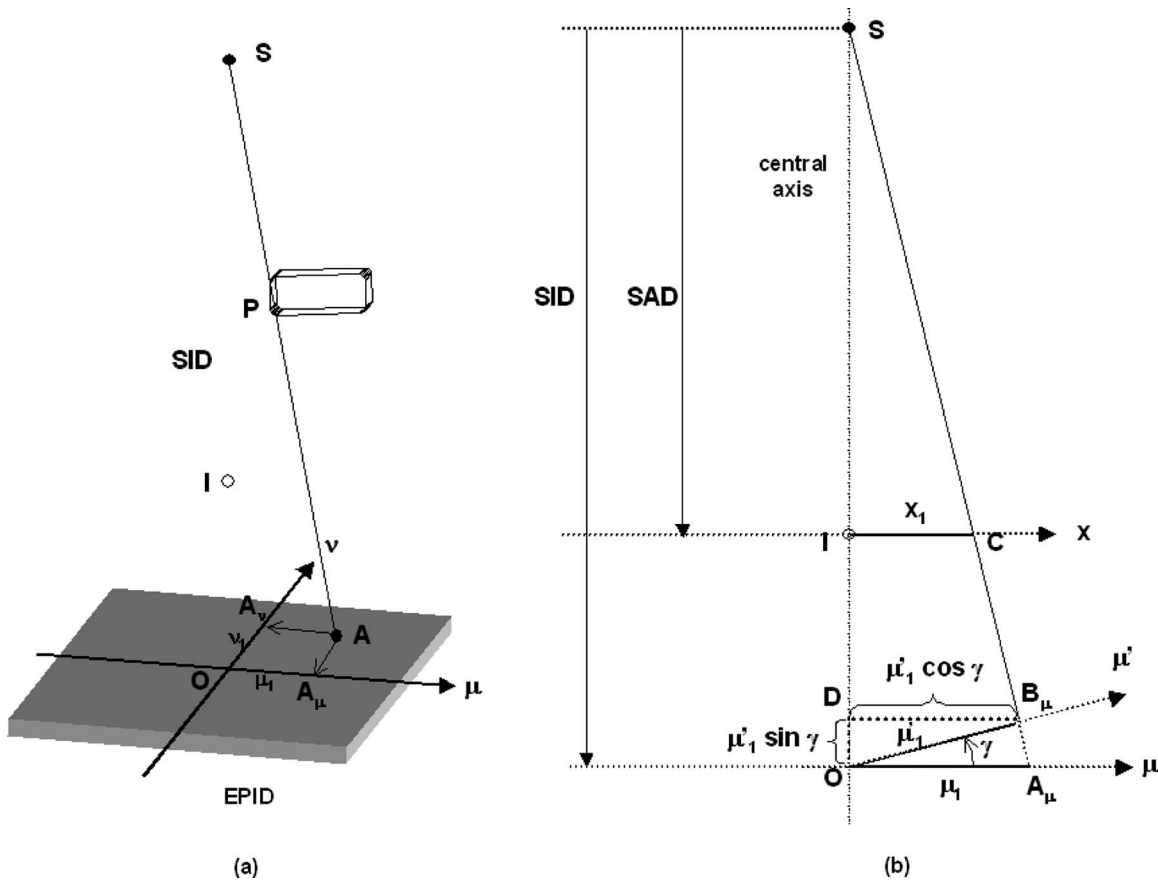


FIG. 2. Schematic diagrams showing the projection of a leaf end P on the EPID (μ, ν) plane (a) and the geometrical relationship of various parameters in the calculation of the leaf end position on the isocentric plane (b). SAD: Source-to-axis distance; SID: Source-to-imager distance.

denotes the isocentric plane with the origin at the isocenter [Fig. 1(a)]. The (μ, ν) plane pertains to the EPID with the origin O at the center of the EPID [Fig. 1(b)]. Both the (x, y) and (μ, ν) planes rotate with the gantry angle ω . The μ' and ν' axes are defined for a tilted EPID; the angles γ and θ represent the tilt of the μ' and ν' axes, respectively [Figs. 1(c) and 1(d)].

Suppose a leaf and P (assumed to be a point) is projected to A on the EPID plane (μ, ν) [Fig. 2(a)]. The projection of A on the μ and ν axes are A_μ and A_ν , respectively. The distances of A_μ and A_ν are at μ_1 and ν_1 from O , respectively. Now consider a vertical plane passing through the x-ray source (S), isocenter (I), and the point A_μ [Fig. 2(b)], and if we assume the imager is in perfect horizontal alignment; the position of the leaf end on the x axis of the isocentric plane can be calculated.

From similar triangles SIC and SOA_μ ,

$$x_1 = \mu_1 \times \frac{SAD}{SID}. \quad (1)$$

Similarly,

$$y_1 = \nu_1 \times \frac{SAD}{SID}. \quad (2)$$

Now assume the imager is tilted to an angle γ , the leaf end is projected to a point B on the tilted EPID instead of A on the

horizontal EPID. The projection of B on the μ axis is B_μ , which is at a distance of μ'_1 from O .

From similar triangles SDB_μ and SOA_μ , the position of the leaf end (μ_1) on the μ axis can be found by

$$\mu_1 = \mu'_1 \cos \gamma \times \frac{SID}{SID - \mu'_1 \sin \gamma}. \quad (3)$$

Using Eq. (1), the corresponding leaf end position on the x axis (x_1) of the isocentric plane is

$$x_1 = \mu'_1 \cos \gamma \times \frac{SAD}{SID - \mu'_1 \sin \gamma}. \quad (4)$$

And if we further incorporate the EPID offset distance ($d_{\mu'}$) along the μ' axis into Eq. (4), we have

$$x_1 = (\mu'_1 + d_{\mu'}) \cos \gamma \times \frac{SAD}{SID - (\mu'_1 + d_{\mu'}) \sin \gamma}. \quad (5)$$

Without loss of generality, the corresponding leaf end position on the y axis (y_1) of the isocentric plane is

$$y_1 = (\nu'_1 + d_{\nu'}) \cos \theta \times \frac{SAD}{SID - (\nu'_1 + d_{\nu'}) \sin \theta}. \quad (6)$$

In a perfect imaging geometry where the tilts γ and θ equal to zero ($\mu' = \mu$; $\nu' = \nu$), and there are no offsets of the

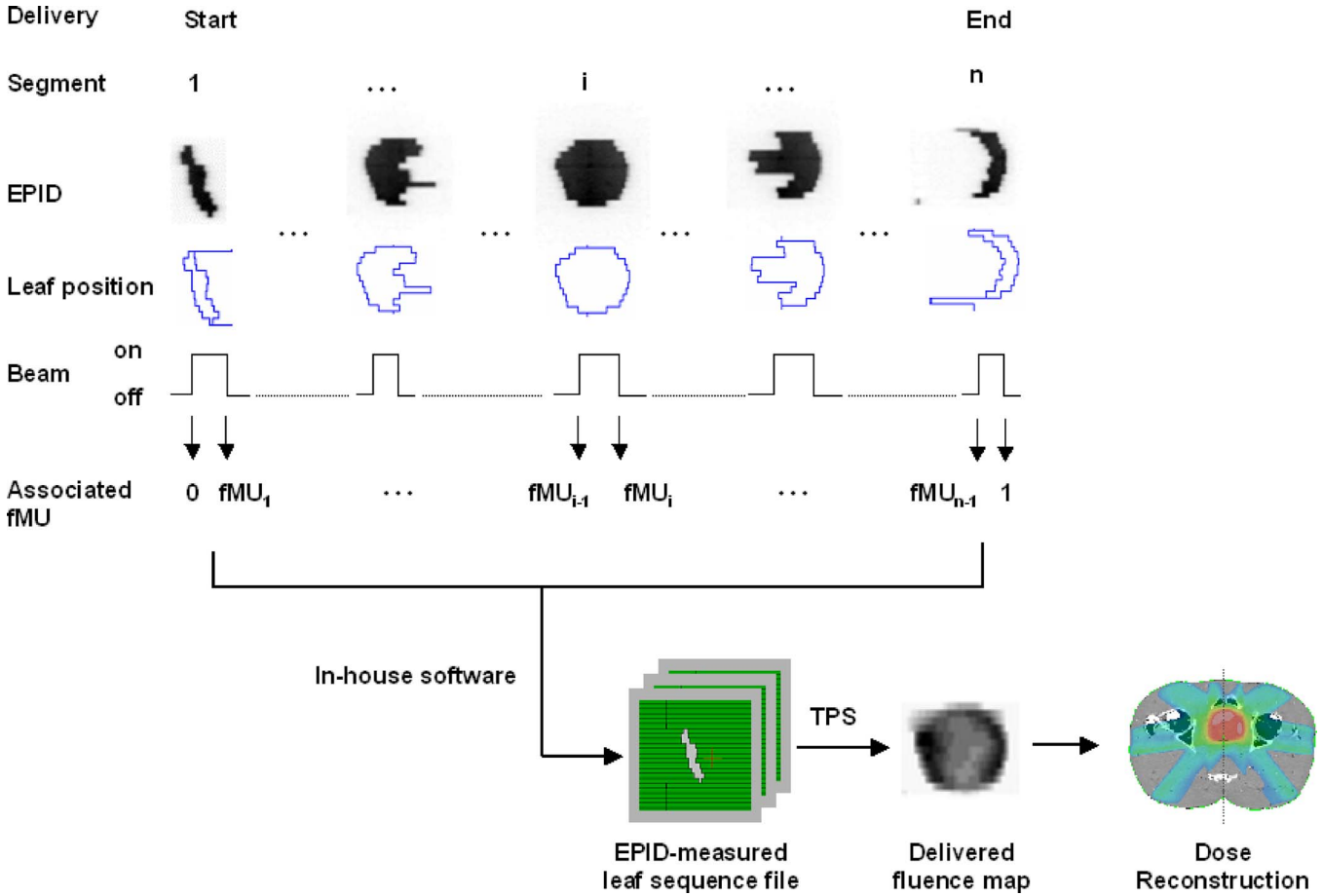


FIG. 3. Schematic diagram showing the workflow from the acquisition of the MLC segments to the dose reconstruction in the prostate IMRT delivery.

EPID center; Eqs. (5) and (6) reduce to Eqs. (1) and (2), respectively.

II.D. Experimental design and software development

A step-and-shoot IMRT field is made up of a number of segments; each segment is specified by prescribed leaf positions and a fraction of the total monitor units to be delivered for this segment. For the experiments, a step-and-shoot IMRT field with an open segment 10 cm long by 1 cm wide (“open gap” field) was designed to step through a distance of 10 cm from −5 to 5 cm on the *x* axis of the isocentric plane.

Five MUs were assigned to each segment. The leaf motion was parallel to the μ axis of the imager. This IMRT field was delivered to a pelvic phantom (with a bony pelvis embedded) with its center aligned to the isocenter at a gantry angle of 0° (IEC scale). The 6 MV beam was used at a dose rate of 300 MU/min. The EPID was positioned at a SID of 150 cm to ensure enough clearance for all gantry angles. During the IMRT delivery, EPID images were acquired in a cine mode at a frame rate of 6.7 frames per second (fps) and the acquisition was synchronized with the beam. Each captured frame was saved as a separate image. There was no user interven-

TABLE I. The mean deviations of the leaves (Banks B and A) for the ten segments of the open gap field delivered at the gantry angle 0° in three repeated deliveries.

		Mean deviation of leaves (mm) from expected positions									
		Segment									
Delivery		1	2	3	4	5	6	7	8	9	10
Bank B	1	0.8	0.7	0.7	0.8	0.9	0.7	0.8	0.8	0.9	0.8
	2	0.8	0.7	0.7	0.8	0.9	0.8	0.7	0.8	0.9	0.8
	3	0.8	0.6	0.7	0.8	0.9	0.7	0.8	0.8	0.9	0.8
Bank A	1	0.8	0.8	0.8	0.8	1.0	0.8	0.8	0.9	0.9	1.0
	1	0.8	0.7	0.8	0.8	1.0	0.9	0.8	1.0	1.0	1.0
	3	0.8	0.7	0.8	0.8	1.0	0.9	0.9	1.0	0.9	0.9

TABLE II. The mean deviations of the leaves (Banks B and A) for the ten segments of the open gap field delivered at the four principal orthogonal gantry angles.

		Mean deviation of leaves (mm) from expected positions									
		Segment									
	Gantry (deg)	1	2	3	4	5	6	7	8	9	10
Bank B	0	0.8	0.7	0.7	0.8	0.9	0.7	0.8	0.8	0.9	0.8
	90	0.4	0.4	0.5	0.5	0.7	0.5	0.5	0.6	0.5	0.6
	180	0.3	0.5	0.5	0.7	0.6	0.6	0.6	0.7	0.7	0.6
	270	0.8	0.7	0.7	1.0	0.9	0.7	0.8	0.8	0.8	0.8
Bank A	0	0.8	0.8	0.8	0.8	1.0	0.8	0.8	0.9	0.9	1.0
	90	0.4	0.4	0.5	0.5	0.7	0.5	0.5	0.5	0.6	0.4
	180	0.4	0.5	0.6	0.7	0.6	0.6	0.7	0.6	0.7	0.7
	270	0.8	0.7	0.8	0.8	0.8	0.7	0.8	0.7	0.8	0.8

tion during the whole acquisition process. In fact, for some segments with small MUs, only two images were captured. Since the acquisition was synchronized with the beam delivery, it was impossible to capture the end of one segment with the start of the next segment. However, due to the ghosting of the EPID, we did see occasionally some overlap of a segment with a very faint residual image from the previous segment, but this did not pose a problem because our software could easily distinguish the two by referring only the maximum gradients encountered. All EPID images were processed with the dark-field and flood-field corrections. The delivery was repeated three times at this angle and also at other orthogonal angles of 90°, 180°, and 270° in order to study the effect of gravity on the measurement of the leaf position from the EPID images.

A further experiment (“open air gap” experiment) was performed in exactly the same settings without the pelvic phantom in the beam in order to see whether the presence of a phantom in the open gap experiment would compromise the detection of the leaf end position due to the extra scatter or bony interface within the captured image.

An in-house program written in MatLab code (MathWorks, Inc., Natick, MA) was developed to reconstitute the EPID-measured LS file. The program first identifies different segments from a series of EPID images captured from the IMRT field by performing morphologic comparison. The same images of a segment are grouped together; the first and last images are taken as the starting and ending shapes for that particular segment. For each segment image, the positions of the leaf ends are searched near the penumbral regions by a maximum gradient edge detection algorithm in a scanline fashion for the pixel rows. The maximum gradient in intensity in the penumbral region of a portal image has been confirmed to correspond to the 50% intensity level^{25,26} that depicts the dosimetric leaf edge.²⁷ The pixel location (column number, row number) of the leaf end found is spatially converted and projected back to the isocentric plane with the positional corrections using Eqs. (5) and (6). Only the pixel rows that correspond to the central one-third of the leaf are calculated to avoid the influence of the interleaf transmission;²⁸ the final leaf end position is defined as an

average of the positions of the leaf ends found from the set of pixel rows that belong to the same leaf.^{19,28,29} The leaf positions of each pair of MLC leaves for each segment image are then coupled with the fMU associated with that particular segment by referring to the synchronized beam on signal and cumulative MU sampled from the beam control circuitry and MLC workstation, respectively.^{25,30} After all the segments in one delivered field have been analyzed, an EPID-measured LS file is reconstituted in a format readable by the TPS using the segmental leaf positions and fMU values. The EPID-measured LS file, which reflects the actual delivery, is then used to derive the delivered fluence map for dose reconstruction.

The measurement of fMU is independent of the frame rate used. The measurement is from the MLC control workstation and console electronics cabinet, whereas the frame rate is an acquisition parameter used in the portal vision. The EPID is only used to capture the leaf position in a segment; the corresponding fMU associated with a particular segment is supplied by the MLC workstation through the coupling of the beam holdoff signal. Since the IMRT delivery is in the step-and-shoot mode and the image acquisition is synchronized with the beam pulses, by referring to the beam holdoff (beam on and off interval), we can associate the cumulative MU sampled from the console electronics cabinet with a particular segment (Fig. 3). In the cine mode, the portal imager actually “waits” for the “beam on” signal to start acquiring the images at 6.7 fps; once the beam is held off during the movement of the leaves from one segment to another, the portal imager will stop acquiring. When the leaves reach their intended positions (within the tolerance) in the next segment, the beam will be activated again and so be the portal imager. This is exactly what we rely on to couple the fMU with a particular segment. All the required data are readily available if one has access to the appropriate ports of the MLC workstation and console electronics cabinet with the help of engineering personnel.

The efficacy of the developed software in measuring the leaf end position was assessed by comparing the measured and expected leaf end positions from the open gap field ex-

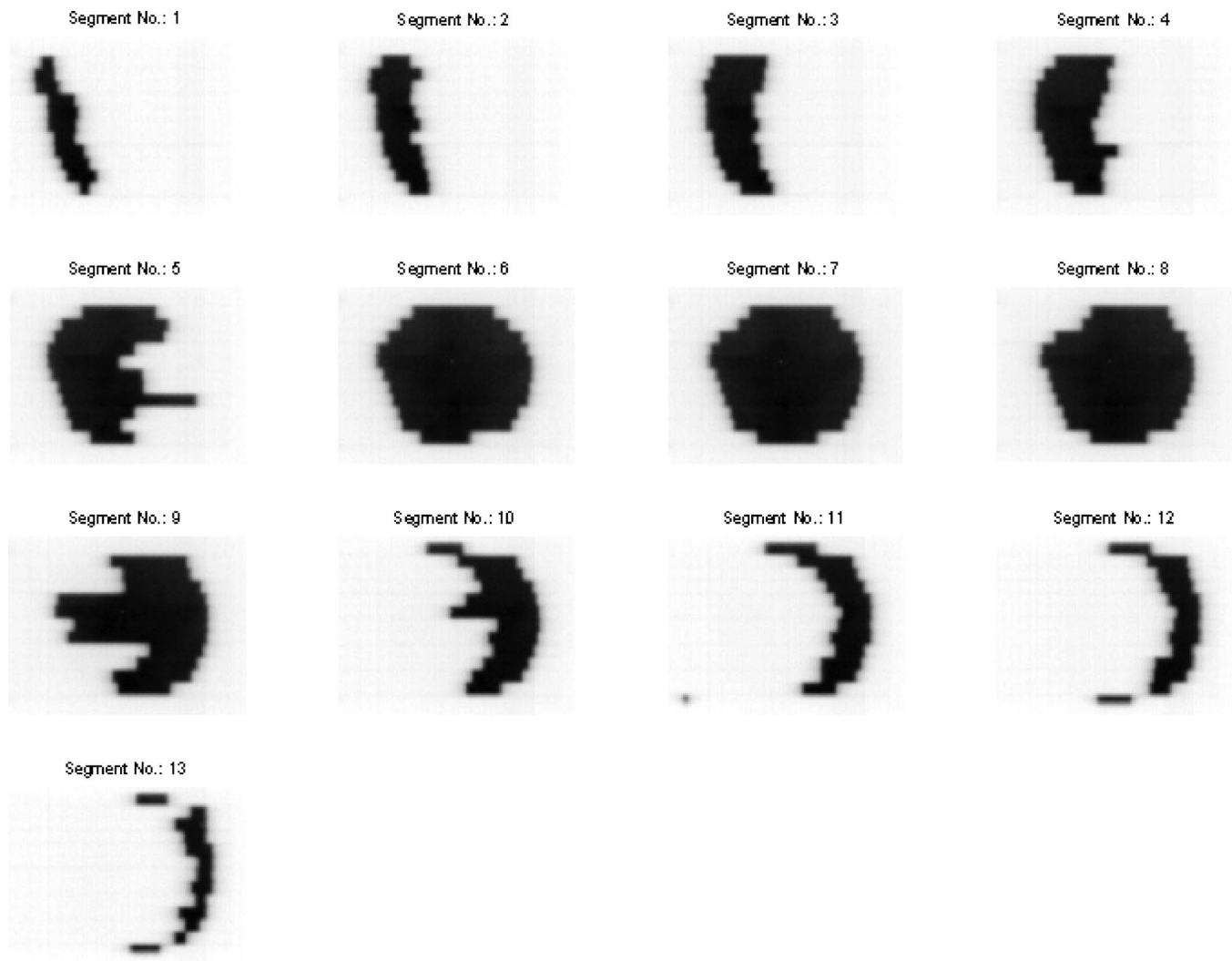


FIG. 4. A series of 13 MLC segments captured by the EPID for the delivered field at the gantry angle 0° for the prostate IMRT delivery.

periments. The differences in the leaf end positions calculated from the open gap and open air gap experiments were also studied.

II.E. Dose reconstruction

A typical prostate step-and-shoot IMRT plan with seven coplanar fields was copied from a patient case and applied to a pelvic phantom. The treatment was delivered to the pelvic phantom as in a real treatment on a Trilogy linac using the 6 MV photon beam at a dose rate of 300 MU/min. EPID images were acquired during the entire delivery as described previously. For each delivered field, there were about 9–13 segments depending on the modulation of the fluence map, and there were about 4–6 EPID images captured for each individual segment, making a total of about 40–80 EPID images for each delivered field. A typical prostate IMRT would result in about 450 EPID images. These images were discarded once the analysis was done to save computer space. The EPID images were analyzed by the developed software, and the reconstituted LS files were loaded back to the ECLIPSE TPS (Varian Medical Systems, Palo Alto, CA) to

derive the delivered fluence maps. The delivered and planned fluence maps were compared. Dose reconstruction was then performed on the pelvic phantom using the delivered fluence maps. The workflow is shown schematically in Fig. 3. Dose reconstruction was also performed using the planned fluence maps. The dose distributions from the two dose reconstructions on the three orthogonal planes through the isocenter and the corresponding dose volume histograms (DVHs) were compared. Note that the plan was not optimized for the pelvic phantom, it is merely used to show the difference in dose reconstruction from using the delivered against planned fluence maps.

III. RESULTS

III.A. Open gap and open air gap field experiments

For each MLC segment in the open gap field, the mean deviation (σ) of the leaves (leaf m to leaf n) forming the gap on each bank from their expected positions is defined as

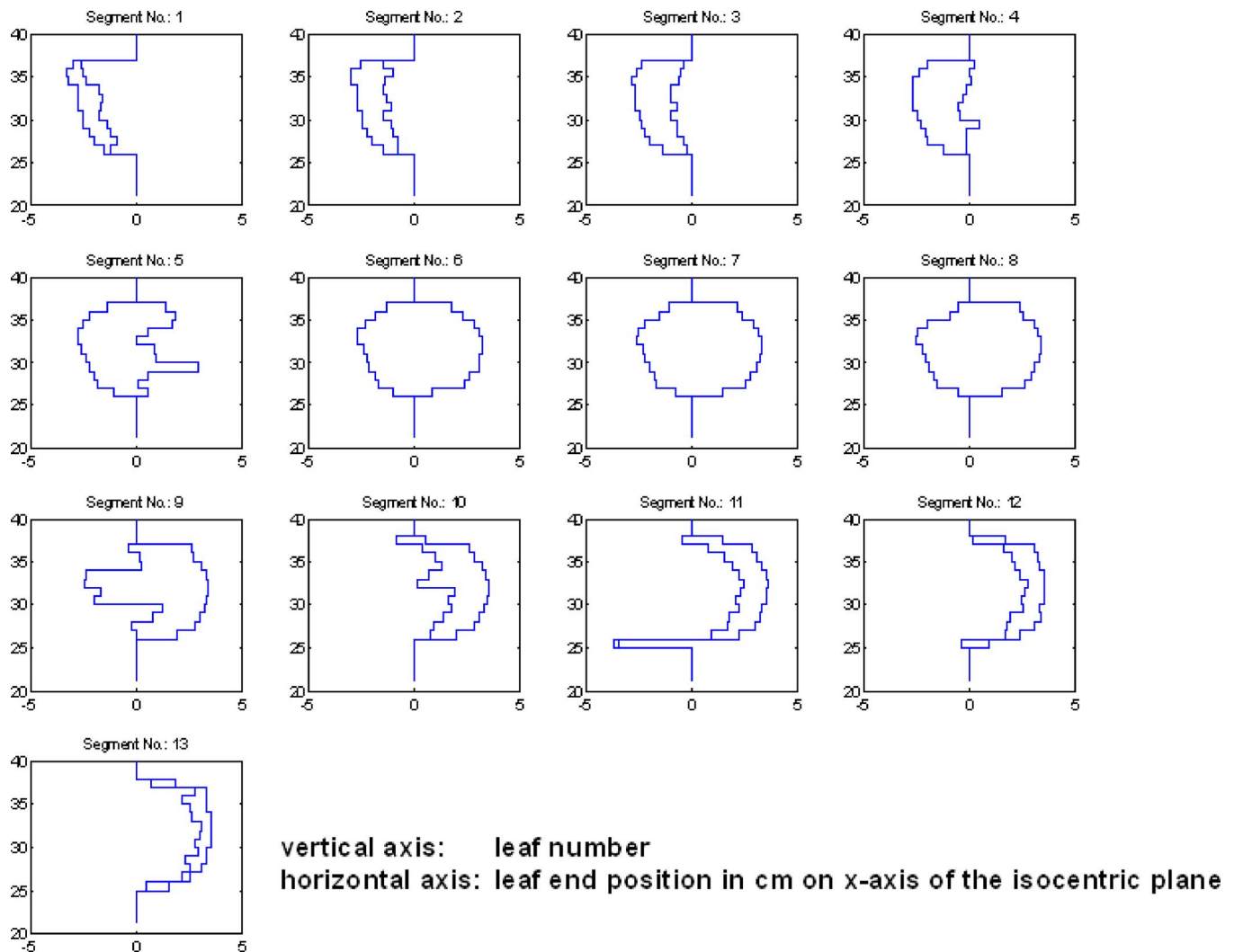


FIG. 5. A series of leaf end positions for each individual MLC segment from Fig. 4.

$$\sigma = \sqrt{\frac{\sum_{i=m}^n (x_i - x_{i,\text{expected}})^2}{n - m + 1}}, \quad (7)$$

where x_i and $x_{i,\text{expected}}$ are the measured and expected leaf end positions of the i th leaf, respectively.

The mean deviations of the leaves for the ten segments (Banks B and A) delivered at the gantry angle 0° were about 0.6–1.0 mm and were reproducible for the three repeated measurements (Table I). The mean deviations of the leaves for the ten segments (Banks B and A) delivered at the four principal orthogonal gantry angles were of the same order of magnitude (Table II). However, we noticed that while the mean deviations for the gantry angle 270° were similar to that at the gantry angle 0° , the mean deviations for the gantry angles of 90° and 180° were found to be smaller (~ 0.3 – 0.7 mm). Similar results were found for the open air gap experiment; data were not shown to avoid redundancy. Concerning these results, we did not find any discernible difference in the leaf end positions whether the phantom was present or not.

III.B. Prostate IMRT delivery: MLC segments captured by EPID

Figure 4 shows a series of 13 segments (only the first image of each segment is shown) captured by the EPID for the delivered field at the gantry angle 0° in the prostate IMRT delivery. The corresponding leaf end positions calculated by the developed software for each individual segment are shown in Fig. 5. The calculated leaf end positions were well within 1.0 mm of their expected positions.

III.C. Derivation of fluence maps and dose reconstruction

The delivered fluence maps derived from the EPID-measured LS files for the seven IMRT fields are shown in Fig. 6(b); there were slight differences (arrows) in the intensity levels when compared to the corresponding planned fluence maps from the original plan [Fig. 6(a)]. Fluence maps directly reconstructed from the MLC log files are also displayed in Fig. 6(c) for comparison. It was found that the

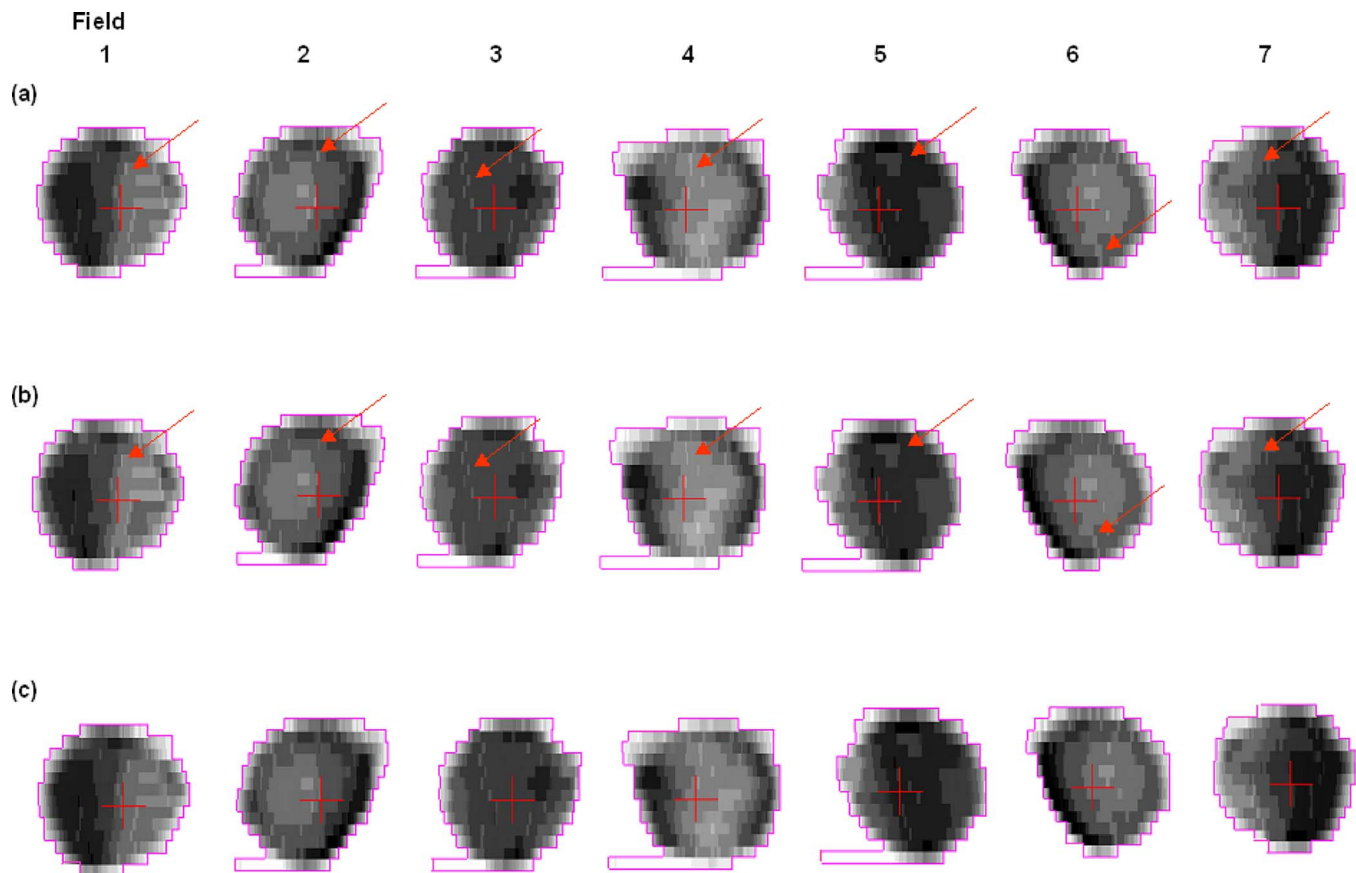


FIG. 6. (a) Planned fluence maps. (b) Delivered fluence maps from EPID-measured leaf sequence files. The differences in the intensity levels between the planned and delivered fluence maps are highlighted by the arrows. No discernible difference is seen between the delivered fluence maps and those directly reconstructed from the MLC log files (c).

fluence maps derived from the EPID-measured LS files could faithfully depict those reconstructed from the MLC log files.

The dose reconstruction using the delivered fluence maps showed no discernible difference for all the isodose lines except the one for the 105% when compared to the expected dose distribution (Fig. 7). The volumes pertaining to a dose level of 105% or above were 12 and 27 cm³ for the delivered and expected dose distributions, respectively. The maximum dose in the delivered dose distribution was 2.5% lower than that in the expected dose distribution. The DVHs of the targets from the two dose reconstructions essentially overlapped with each other except at the high dose regions, which are consistent with the difference seen in the dose distributions.

IV. DISCUSSION

We have shown the feasibility in reconstituting an EPID-measured LS file and using it to derive the delivered fluence map for dose reconstruction. The approach is equivalent to using the MLC log files for dose reconstruction^{11,13} and yet avoids the dependency of the recorded leaf positions on the leaf calibration. The algorithm of detecting the leaf end position is, in principle, independent of the EPID used. The principle of obtaining the dose information can easily be

adopted for linacs of other vendors provided that the console electronics cabinet has an interface to assess the beam hold-off signal and cumulative MU.

At the outset, our main concern was the influence of EPID sag and tilt at different gantry angles on the measurement of the leaf end position. Clark and Budgell²⁰ and Woo *et al.*³⁰ have solved this problem by calibration techniques. Using the developed gQA tool, we were able to quantify the EPID sag and tilt and incorporate this information in the calculation of the leaf end positions.²⁴ The use of the present approach is limited by the projected field size on the EPID, which is 20 × 26 cm² at the isocentric plane.

From the results of the open gap field experiments, we found that the mean deviation of the leaves was about 0.6–1.0 mm, which was expected of a MLC-based IMRT delivery^{28,31} taking into account that (1) a tolerance of 1.5 mm in the leaf position was allowed during the actual delivery, (2) the associated image noise somewhat compromised the accurate measurement of the leaf end at the penumbral region, and (3) the slight effect of leaf sagging at various gantry angles.²⁰ The projected pixel size at the isocentric plane is 0.26 mm, which means that the mean deviation was about 4 pixels. We found that the mean deviations of the leaves at gantry angles of 90° and 180° (~0.6 mm) were in general smaller than that at the gantry angles 0° and

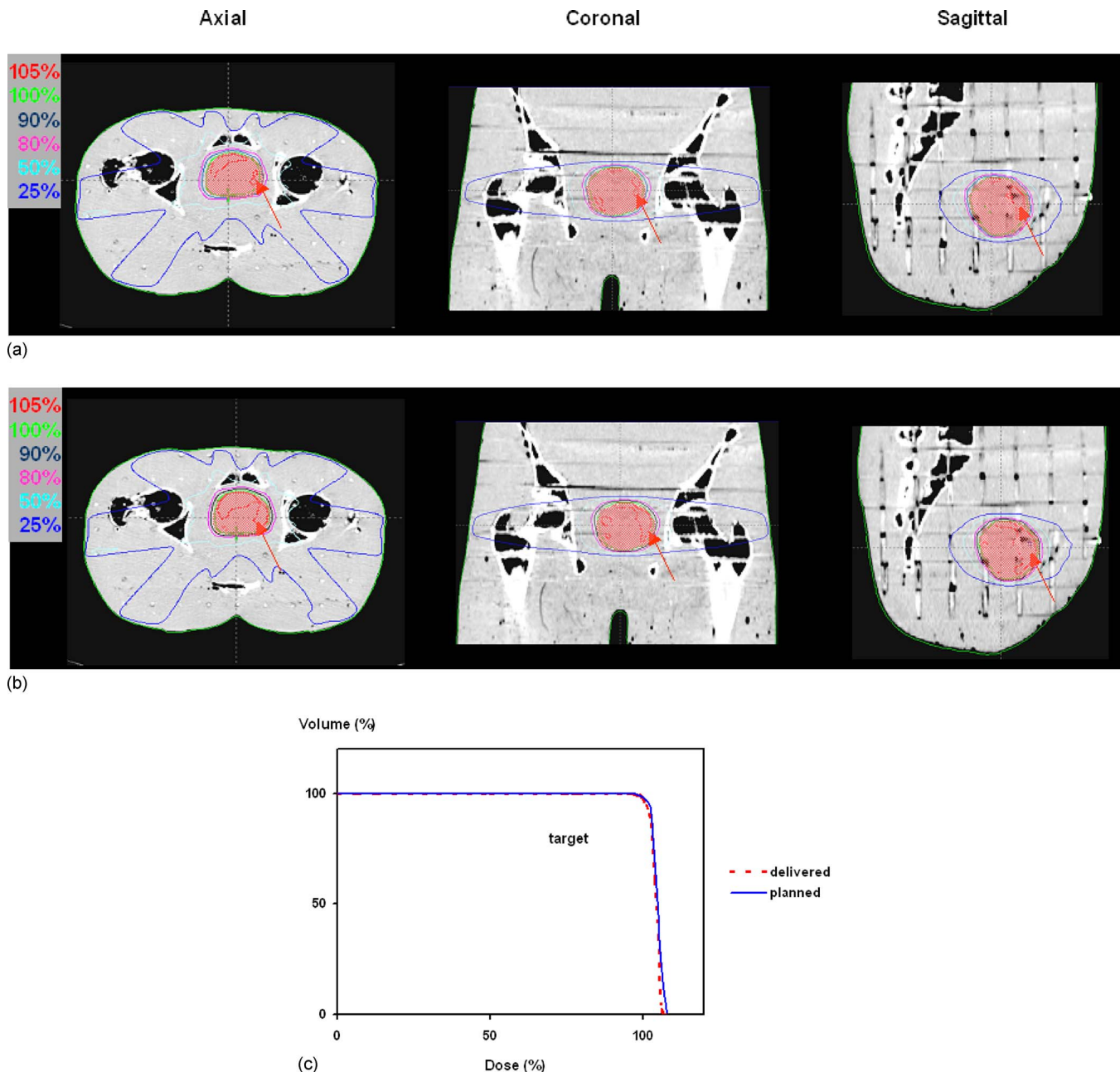
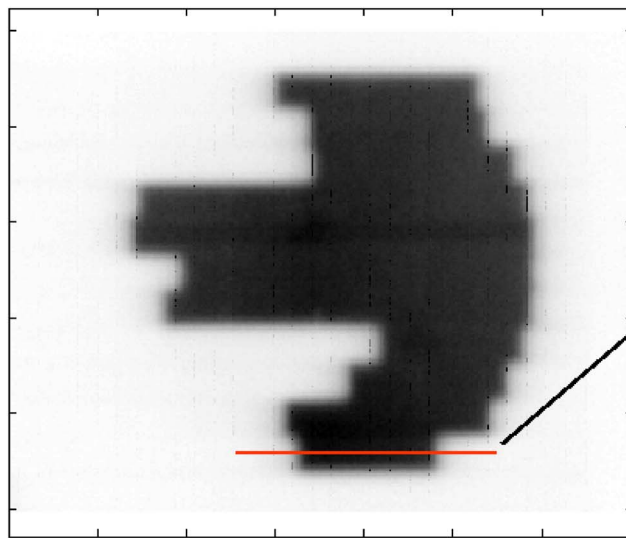


FIG. 7. (a) Delivered dose distribution calculated with fluence maps derived from EPID-measured leaf sequence files. (b) Expected dose distribution calculated with planned fluence maps. The differences in the dose distribution (105% level) between the delivered and expected dose distributions for the three orthogonal planes are highlighted by the arrows. (c) The DVHs of the targets from the two dose reconstructions essentially overlapped each other except at the dose level beyond 105%.

270°, and there was no discernible difference in the mean deviations for the two different banks when the gantry was horizontally placed (gantry angles of 90° and 270°). Under the present experimental setup, the data we acquired were insufficient to explain this phenomenon; one possible explanation could be that the routine leaf calibration is done at the gantry angle 0° and it might deviate from the ideal scenario at other gantry angles. Moreover, a smaller mean deviation does not necessarily mean it is more accurate in the leaf position at these gantry angles; it simply means the condition of the leaf positioning was different from that at gantry angles of 0° and 180°.

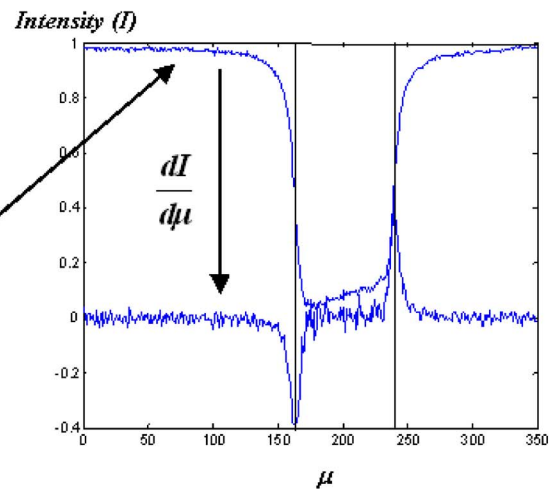
This work is designed to capture the segments in real time with a patient in the beam during the actual delivery. The predeposited condition is that a patient is always present; in fact, the presence of the patient (or the phantom) does not compromise the detection of the leaf ends in our algorithm using the maximum gradient search. The edge detection algorithm is applicable in open fields or fields with patient anatomy. The main reason being the poor image contrast in the MV energy range; even at the bone–soft tissue interface, the gradient is much smaller than at the penumbral region where there is a maximum transition from a region of background noise to an irradiated region. The poor contrast also

pixel no. along v -direction



pixel no. along μ -direction

(a)



- leaf end position = 50% intensity level
- @ maximum gradient

(b)

FIG. 8. (a) Due to the poor image contrast in MV energy range, the bony structures (present in the pelvic phantom) are hardly seen in the image. (b) The gradients encountered at the leaf ends overwhelm all other gradients that would otherwise be present due to any bony interfaces.

explains why the bony details are not seen in the EPID images. An EPID image is shown in Fig. 8 to demonstrate the poor contrast and the magnitude of the order of the gradient encountered at the field edge and within the field where there was actually some bony anatomy.

In the comparison of the delivered fluence maps with the planned ones, the differences in the intensity levels seen were attributable to the combined effect of leaf deviations in each of the contributing segments and the inherent fMU redistribution;^{8,13} the dosimetric impact of these differences was reflected in the dose reconstruction using the delivered fluence maps. The differences in the 105% isodose level and the maximum dose between the delivered and expected dose distributions indicate that a need for employing the delivered rather than planned fluence maps for dose reconstruction in ART is warranted.

Unfortunately, the present technique does not work for dynamic delivery (sliding window). Since we rely on the beam holdoff signal while the leaves are stepping across to determine the fMU association with each segment, we have not yet figured out how we can associate a fMU with an ever-changing segment in the sliding window delivery.

One might argue that dose reconstruction should be accomplished by *in vivo* exit portal dosimetry,^{32,33} but the technique is still in its development and impeded by some confounding issues such as dosimetric nonlinearities of the aSi-EPID^{17,34} that require tedious calibration and correction³⁵ and the correction of scatter from the patient.³² Our approach

is developed along the line of using MLC log files for dose reconstruction and based on a well-proven superiority of aSi-EPID in spatial determination of the leaf end position;^{14,19–21} it is straightforward and easy to implement at a clinic. In addition, the proposed approach is equally well suited for IMRT dose verification and quality assurance purposes.

V. CONCLUSIONS

This paper presented a proof of the feasibility and principles in reconstituting a leaf sequence file that reflects the actual delivery of an IMRT based on the MLC segments captured by an EPID. This EPID-measured LS file can serve the same purpose as a MLC log file does for the derivation of the delivered fluence map and is independent of the leaf calibration. The approach also allows users who do not have access to MLC log files to probe the actual IMRT delivery and translate the information gained for dose reconstruction in ART.

ACKNOWLEDGMENTS

Support from the Sir Robert Black Postdoctoral Fellowship and the Hong Kong Hospital Authority Overseas Training Allowance for the first author are gratefully acknowledged. This work is also supported in part by grants from the Varian Medical Systems, Department of Defense (No. W81XWH-05-1-0041) and National Cancer Institute (No. 5R01 CA98523). The authors would also like to thank Mark

Wanlass and Bill Cheng, our linac engineers, for offering technical support and advice on the control of the linac and MLC.

- ^{a)} Author to whom correspondence should be addressed. Telephone: (650) 498-7896; Fax: (650) 498-4015. Electronic mail: lei@reyes.stanford.edu
- ¹ S. Webb, "Adapting IMRT delivery fraction-by-fraction to cater for variable intrafraction motion," *Phys. Med. Biol.* **53**, 1–21 (2008).
 - ² A. de la Zerda, B. Armbruster, and L. Xing, "Formulating adaptive radiation therapy (ART) treatment planning into a closed-loop control framework," *Phys. Med. Biol.* **52**, 4137–4153 (2007).
 - ³ D. Yan, F. Vicini, J. Wong, and A. Martinez, "Adaptive radiation therapy," *Phys. Med. Biol.* **42**, 123–132 (1997).
 - ⁴ D. Yan, J. Wong, F. Vicini, J. Michalski, C. Pan, A. Frazier, E. Horwitz, and A. Martinez, "Adaptive modification of treatment planning to minimize the deleterious effects of treatment setup errors," *Int. J. Radiat. Oncol., Biol., Phys.* **38**, 197–206 (1997).
 - ⁵ S. Yoo and F. F. Yin, "Dosimetric feasibility of cone-beam CT-based treatment planning compared to CT-based treatment planning," *Int. J. Radiat. Oncol., Biol., Phys.* **66**, 1553–1561 (2006).
 - ⁶ Y. Yang, E. Schreibmann, T. Li, C. Wang, and L. Xing, "Evaluation of on-board kV cone beam CT (CBCT)-based dose calculation," *Phys. Med. Biol.* **52**, 685–705 (2007).
 - ⁷ K. M. Langen, S. L. Meeks, D. O. Poole, T. H. Wagner, T. R. Willoughby, P. A. Kupelian, K. J. Ruchala, J. Haimerl, and G. H. Olivera, "The use of megavoltage CT (MVCT) images for dose recomputations," *Phys. Med. Biol.* **50**, 4259–4276 (2005).
 - ⁸ P. Xia, C. F. Chuang, and L. J. Verhey, "Communication and sampling rate limitations in IMRT delivery with a dynamic multileaf collimator system," *Med. Phys.* **29**, 412–423 (2002).
 - ⁹ R. D. Wiersma and L. Xing, "Examination of geometric and dosimetric accuracies of gated step-and-shoot intensity modulated radiation therapy," *Med. Phys.* **34**, 3962–3970 (2007).
 - ¹⁰ D. W. Litzenberg, S. W. Hadley, N. Tyagi, J. M. Balter, R. K. Ten Haken, and I. J. Chetty, "Synchronized dynamic dose reconstruction," *Med. Phys.* **34**, 91–102 (2007).
 - ¹¹ L. Lee, Q. T. Le, and L. Xing, "Retrospective IMRT dose reconstruction based on cone-beam CT and MLC log-file," *Int. J. Radiat. Oncol., Biol., Phys.* **70**, 634–644 (2008).
 - ¹² J. G. Li, J. F. Dempsey, L. Ding, C. Liu, and J. R. Palta, "Validation of dynamic MLC-controller log files using a two-dimensional diode array," *Med. Phys.* **30**, 799–805 (2003).
 - ¹³ A. M. Stell, J. G. Li, O. A. Zeidan, and J. F. Dempsey, "An extensive log-file analysis of step-and-shoot intensity modulated radiation therapy segment delivery errors," *Med. Phys.* **31**, 1593–1602 (2004).
 - ¹⁴ L. Parent, J. Seco, P. M. Evans, D. R. Dance, and A. Fielding, "Evaluation of two methods of predicting MLC leaf positions using EPID measurements," *Med. Phys.* **33**, 3174–3182 (2006).
 - ¹⁵ G. V. Menon and R. S. Sloboda, "Quality assurance measurements of a-Si EPID performance," *Med. Dosim.* **29**, 11–17 (2004).
 - ¹⁶ B. M. McCurdy, K. Luchka, and S. Pistorius, "Dosimetric investigation and portal dose image prediction using an amorphous silicon electronic portal imaging device," *Med. Phys.* **28**, 911–924 (2001).
 - ¹⁷ G. J. Budgell, Q. Zhang, R. J. Troncner, and R. I. Mackay, "Improving IMRT quality control efficiency using an amorphous silicon electronic portal imager," *Med. Phys.* **32**, 3267–3278 (2005).
 - ¹⁸ J. Chang, C. H. Obcemea, J. Sillanpaa, J. Mechalakos, and C. Burman, "Use of EPID for leaf position accuracy QA of dynamic multi-leaf collimator (DMLC) treatment," *Med. Phys.* **31**, 2091–2096 (2004).
 - ¹⁹ S. J. Baker, G. J. Budgell, and R. I. MacKay, "Use of an amorphous silicon electronic portal imaging device for multileaf collimator quality control and calibration," *Phys. Med. Biol.* **50**, 1377–1392 (2005).
 - ²⁰ M. F. Clarke and G. J. Budgell, "Use of an amorphous silicon EPID for measuring MLC calibration at varying gantry angle," *Phys. Med. Biol.* **53**, 473–485 (2008).
 - ²¹ J. J. Sonke, L. S. Ploeger, B. Brand, M. H. Smitsmans, and M. van Herk, "Leaf trajectory verification during dynamic intensity modulated radiotherapy using an amorphous silicon flat panel imager," *Med. Phys.* **31**, 389–395 (2004).
 - ²² M. N. Graves, A. V. Thompson, M. K. Martel, D. L. McShan, and B. A. Fraass, "Calibration and quality assurance for rounded leaf-end MLC systems," *Med. Phys.* **28**, 2227–2233 (2001).
 - ²³ Y. Yang and L. Xing, "Quantitative measurement of MLC leaf displacements using an electronic portal image device," *Phys. Med. Biol.* **49**, 1521–1533 (2004).
 - ²⁴ W. Mao, L. Lee, and L. Xing, "Development of a QA phantom and automated analysis tool for geometric quality assurance of on-board MV and kV x-ray imaging systems," *Med. Phys.* **35**(4), 1497–1506 (2008).
 - ²⁵ M. Partridge, P. M. Evans, A. Mosleh-Shirazi, and D. Convery, "Independent verification using portal imaging of intensity-modulated beam delivery by the dynamic MLC technique," *Med. Phys.* **25**, 1872–1879 (1998).
 - ²⁶ J. Bijhold, K. G. Gilhuijs, M. van Herk, and H. Meertens, "Radiation field edge detection in portal images," *Phys. Med. Biol.* **36**, 1705–1710 (1991).
 - ²⁷ ICRU, Report No. 24, 1976 (unpublished).
 - ²⁸ S. C. Vieira, M. L. Dirks, K. L. Pasma, and B. J. Heijmen, "Fast and accurate leaf verification for dynamic multileaf collimation using an electronic portal imaging device," *Med. Phys.* **29**, 2034–2040 (2002).
 - ²⁹ H. V. James, S. Atherton, G. J. Budgell, M. C. Kirby, and P. C. Williams, "Verification of dynamic multileaf collimation using an electronic portal imaging device," *Phys. Med. Biol.* **45**, 495–509 (2000).
 - ³⁰ M. K. Woo, A. W. Lightstone, G. Shan, L. Kumaraswamy, and Y. Li, "Automatic verification of step-and-shoot IMRT field segments using portal imaging," *Med. Phys.* **30**, 348–351 (2003).
 - ³¹ K. Eilertsen, "Automatic detection of single MLC leaf positions with corrections for penumbral effects and portal imager dose rate characteristics," *Phys. Med. Biol.* **42**, 313–334 (1997).
 - ³² L. N. McDermott, M. Wendling, J. Nijkamp, A. Mans, J. J. Sonke, B. J. Mijnheer, and M. van Herk, "3D in vivo dose verification of entire hypofractionated IMRT treatments using an EPID and cone-beam CT," *Radiother. Oncol.* **86**, 35–42 (2008).
 - ³³ W. J. van Elmpt, S. M. Nijsten, A. L. Dekker, B. J. Mijnheer, and P. Lambin, "Treatment verification in the presence of inhomogeneities using EPID-based three-dimensional dose reconstruction," *Med. Phys.* **34**, 2816–2826 (2007).
 - ³⁴ L. N. McDermott, S. M. Nijsten, J. J. Sonke, M. Partridge, M. van Herk, and B. J. Mijnheer, "Comparison of ghosting effects for three commercial a-Si EPIDs," *Med. Phys.* **33**, 2448–2451 (2006).
 - ³⁵ S. M. Nijsten, W. J. van Elmpt, M. Jacobs, B. J. Mijnheer, A. L. Dekker, P. Lambin, and A. W. Minken, "A global calibration model for a-Si EPIDs used for transit dosimetry," *Med. Phys.* **34**, 3872–3884 (2007).

Using total-variation regularization for intensity modulated radiation therapy inverse planning with field-specific numbers of segments

Lei Zhu¹, Louis Lee¹, Yunzhi Ma¹, Yinyu Ye², Rafe Mazzeo³
and Lei Xing¹

¹ Department of Radiation Oncology, Stanford University, Stanford, CA 94305, USA

² Department of Management Science and Engineering, Stanford University, Stanford, CA 94305, USA

³ Department of Mathematics, Stanford University, Stanford, CA 94305, USA

E-mail: leizhu@stanford.edu

Received 21 July 2008, in final form 5 September 2008

Published 7 November 2008

Online at stacks.iop.org/PMB/53/6653

Abstract

Currently, there are two types of treatment planning algorithms for intensity modulated radiation therapy (IMRT). The beamlet-based algorithm generates beamlet intensity maps with high complexity, resulting in large numbers of segments in the delivery after a leaf-sequencing algorithm is applied. The segment-based direct aperture optimization (DAO) algorithm includes the physical constraints of the deliverable apertures in the calculation, and achieves a conformal dose distribution using a small number of segments. However, the number of segments is pre-fixed in most of the DAO approaches, and the typical random search scheme in the optimization is computationally intensive. A regularization-based algorithm is proposed to overcome the drawbacks of the DAO method. Instead of smoothing the beamlet intensity maps as in many existing methods, we include a total-variation term in the optimization objective function to reduce the number of signal levels of the beam intensity maps. An aperture rectification algorithm is then applied to generate a significantly reduced number of deliverable apertures. As compared to the DAO algorithm, our method has an efficient form of quadratic optimization, with an additional advantage of optimizing field-specific numbers of segments based on the modulation complexity. The proposed approach is evaluated using two clinical cases. Under the condition that the clinical acceptance criteria of the treatment plan are satisfied, for the prostate patient, the total number of segments for five fields is reduced from 61 using the Eclipse planning system to 35 using the proposed algorithm; for the head and neck patient, the total number of segments for seven fields is reduced from 107 to 28. The head and neck result is also compared to that using an equal number of four segments for each field.

The comparison shows that using field-specific numbers of segments achieves a much improved dose distribution.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Traditional inverse planning algorithms for step-and-shoot IMRT divide the beam's eye view (BEV) of the planning target volume (PTV) into small beamlets (Ezzell *et al* 2003, Xing *et al* 2005). The beamlet intensities are first optimized, and a leaf-sequencing algorithm is then applied on the resultant beam intensity map to generate a set of deliverable beam segments (also called apertures) (Bortfeld *et al* 1994, Xia and Verhey 1998, Kamath *et al* 2004, Saw *et al* 2001, Kuterden and Cho 2001, Ma *et al* 1999). For a fast calculation, the optimization usually has a convex objective function and is typically solved by linear or quadratic programming. Since the physical constraints of the multileaf collimator (MLC) are not included in the optimization, these algorithms usually result in a sub-optimal treatment plan with a large number of beam segments. To reduce the field complexity and therefore the number of beam segments, many algorithms have been proposed in the literature using smoothing techniques (Alber and Nsslin 2000, Ma 2002, Spirou *et al* 2001, Webb *et al* 1998, Sun and Xia 2004, Xiao *et al* 2004). Typical examples use an additional term of sum of derivative squares (Alber and Nsslin 2000, Matuszak *et al* 2007, Spirou *et al* 2001), which are often referred to as quadratic smoothing or regularization in the theory of convex optimization. Since these algorithms smooth the sharp edges of the intensity field as well, the optimized beam intensity is not piecewise constant and still cannot be delivered using a very small number of segments. Another category of optimization algorithms are direct aperture optimization (DAO) methods which have been proposed to eliminate the need for leaf sequencing and include the MLC hardware constraints in the optimization (Shepard *et al* 2002, Michalski *et al* 2004, Cotrutz and Xing 2003, van Asselen *et al* 2006, Bedford and Webb 2007, Bergman *et al* 2006, Mestrovic *et al* 2007). However, because the delivered dose depends on the aperture shapes nonlinearly and the optimization problem is non-convex, random search algorithms, such as simulated annealing, are commonly employed. The computation is intensive and requires tuning of multiple algorithm parameters in the searching and cooling schedules. The issue is exacerbated in advanced applications such as 4D and adaptive therapy treatment planning. Another disadvantage of most of the DAO methods is that the number of segments for each field needs to be determined before the calculation, which reduces the degree of freedom of the decision variable space and compromises the optimality of the final solution. Some sophisticated methods have been proposed to find an optimal set of apertures using floating numbers of segments in the optimization (Romeijn *et al* 2005).

In this work, we propose an efficient inverse planning algorithm which significantly reduces the total number of segments without degrading the conformity of the delivered dose distribution. This algorithm is also able to optimize the number of segments for each field based on the modulation complexity. Instead of directly applying the physical constraints of the apertures, we include a total-variation regularization in the least square optimization to enforce the computed field intensity maps close to be deliverable using a small number of segments. An aperture rectification algorithm is then applied on the intermediate field intensity maps to generate actual intensity maps. The proposed algorithm requires a simple

adaptation of the traditional beamlet-based optimization. Different from other regularization-based methods using smoothing techniques, our method uses the total-variation regularization to reduce the field complexity as well as to encourage a piecewise constant solution (Rudin *et al* 1992). As a result, the number of optimized beam segments is much smaller than that obtained using smoothing methods.

2. Method

2.1. Formulation of the problem

The goal of the inverse treatment planning of radiation therapy is to optimize the treatment beam apertures and weights such that the delivered dose on the patient best matches the prescribed dose. The delivered dose distribution on the patient, d , has a linear relationship with the intensity of the beamlets, x

$$d = Ax, \quad (1)$$

where the 3D dose distribution, d , is vectorized, and the beamlet intensity x is a 1D vector that consists of rowwise concatenations of beamlet intensities for all fields. Each column of matrix A is a beamlet kernel, corresponding to the dose distribution achieved by one beamlet with unit intensity. The beamlet kernels are pre-computed based on the CT image of the patient, the treatment machine settings and the beam geometry. In this work, we used the voxel-based Monte Carlo algorithm (VMC) as our dose calculation engine (Kawrakow *et al* 1996, Kawrakow 1997). The gantry geometry and field angles were based on the typical field setup of the Varian Linac treatment machine. Although matrix A typically has a large size, it is sparse and reduction of computation cost and memory usage is possible (Cho and Phillips 2001, Thieke *et al* 2002).

If the sum of the square errors of the delivered dose relative to the the prescribed dose is used as the optimization objective function, the treatment planning problem can be expressed as

minimize

$$\sum_i \lambda_i (A_i x - d_i)^T (A_i x - d_i) \quad (2)$$

subject to

$$x \geq 0$$

x is achievable using apertures (aperture constraint).

where the index i denotes different structures; λ_i is the relative importance factor (Xing *et al* 1999, Breedveld *et al* 2006, Bortfeld 1999, Webb *et al* 1998, Oelfke and Bortfeld 1999); each column of matrix A_i is the beamlet kernel corresponding to the i th structure, and d_i is the prescribed dose. Note that we use the same form of objective functions for both the target and the critical structures. A low prescribed dose is used for the critical structures. The aperture constraint stems from the physical constraints of the MLC. Specifically, it requires that for each segment, the nonzero beamlets have the same intensity and they are connected in the direction of MLC leaf pairs.

Table 1. Variable glossary.

$A(A_i)$	Matrix that relates the beamlet intensity to delivered dose
$d(d_i)$	Delivered dose
N	Total number of beamlets, $N = N_u N_v N_f$
N_f	Number of fields
N_t	Total number of segments (apertures)
N_u	Number of MLC leaf positions for each leaf
N_v	Number of MLC leaf pairs per field
x	Beamlet intensity, the decision variable in the optimization
λ_i	Importance factor associated with the i th structure
β	Penalty weight associated with the term of total variation

The main variables used in this paper are summarized in table 1 for readers' reference.

2.2. Treatment planning using total-variation regularization

The formulation of the treatment planning problem described above is commonly used in the literature. However, the optimization is not readily implementable. A mathematical expression is still needed for the aperture constraint. Unfortunately, this constraint is neither linear nor quadratic, resulting in a complicated optimization problem. In the conventional treatment planning, the aperture constraint is ignored, and the problem is reduced to a beamlet-based optimization, which can be easily solved as a linear least square problem. In a second step, a leaf-sequencing algorithm is carried out to find the best matched segments for the optimal beamlet intensities. In the DAO algorithms, the aperture constraint is enforced in the choice of the decision variables x in every step of the search process. Due to the nonlinearity of the constraint, a complicated search scheme, such as simulated annealing, is typically used.

In this work, a regularization-based algorithm is proposed to enforce the aperture constraint in the optimization. Instead of smoothing the beamlet intensity maps as in many existing methods, we include a total-variation term in the optimization objective function to encourage a piecewise constant solution and therefore to reduce the number of signal levels of the beamlet intensity maps. The optimization problem is still solved using quadratic programming. The resulting optimized beamlet intensity has sharp spatial transitions. An aperture rectification algorithm is then applied on the optimized beamlet intensity maps to derive a small number of deliverable apertures. Although the aperture rectification algorithm functions similarly to the conventional leaf-sequencing algorithms, it is derived differently with considerations of the unique features of the optimized intensity maps using total-variation regularization.

2.2.1. Beamlet intensity optimization using total-variation regularization. A piecewise constant fluence map, or a fluence map with a small number of signal levels is desirable for reducing the number of segments in the IMRT delivery. The total-variation regularization as used in many optimization problems is able to force the optimized solution to be piecewise constant (Block *et al* 2007, Kolehmainen *et al* 2006). Inspired by these facts, in our new optimization algorithm, we include in the objective function an additional term of total variation (L-1 norm) of the beam intensity x , and the aperture constraint in the optimization problem (2) is removed. The term of total variation calculates the sum of absolute values of the derivatives, and the penalties drive the derivatives toward zeros and force the optimized beam intensity to be close to piecewise constant. We reformulate the optimization problem as

minimize

$$\sum_{i=1}^N \lambda_i (A_i x - d_i)^T (A_i x - d_i) + \beta \sum_{f=1}^{N_f} \sum_{u=2}^{N_u} \sum_{v=2}^{N_v} (|x_{u,v,f} - x_{u-1,v,f}| + |x_{u,v,f} - x_{u,v-1,f}|) \quad (3)$$

subject to

$$x \succeq 0.$$

In the second term of the objective function, the beamlet intensity x is parameterized by the variables u , v and f . The variable u (v) is the row (column) index of the beam intensity for each field and f is the field index. N_u is the total number of possible MLC leaf positions for each leaf; N_v is the total number of MLC leaf pairs per field; N_f is the number of fields. For simplicity, we assume that each treatment field has a rectangular shape when it is fully open, and N_u and N_v do not change for different fields. The parameter β is the penalty weight associated with the term of total variation. Since the deliverable fluence map has sharp transitions only in the horizontal and vertical directions, we calculate the term of total variation as a summation of the absolute values of the one-dimensional derivatives of each beamlet in the u and v directions, instead of two-dimensional derivatives.

The term of total variation does not appear to be linear or quadratic. However, the above optimization can be easily reformulated as a quadratic programming problem (see appendix A). The problem is then efficiently solved using standard quadratic programming.

Although not exactly equivalent, the total-variation regularization in the optimization implies the aperture constraint on the intensity maps. The resulting optimized intensity maps are more readily achievable using apertures, at the price of slightly degraded dose distribution as compared to that of the beamlet-based plan. This tradeoff can be adjusted by using different values of β . If the number of segments is large (small), the complexity of deliverable intensity maps can be high (low) and a small (large) β should be used to relax (strengthen) the aperture constraint. For a given number of segments, too large a β forces the resulting intensity maps deliverable using a very small number of segments, and the capability of high complexity of deliverable intensity maps is not fully utilized; if β is too small, the optimization is close to the conventional beamlet-based method, which requires a number of segments much larger than the given value. In this work, we choose the value of β empirically based on clinical data. More discussion on the choice of β is included in the section of results.

2.2.2. Aperture rectification algorithm. The optimization with the total-variation regularization (3) is not equivalent to the initial optimization problem (2). This is because the aperture constraint is not equivalent to the constraint of piecewise constant beamlet intensity maps, not to mention that including the term of total variation in the objective function does not guarantee the optimized variable to be exactly piecewise constant. A leaf-sequencing algorithm is needed to modify the optimized result of the beamlet intensity maps to enforce the aperture constraint. In this paper, based on the fact that the total-variation regularization greatly reduces the number of signal levels of the optimized intensity maps, we develop a novel aperture rectification algorithm for the leaf sequencing. The details of the algorithm along with an implementation example are included in appendix B. Note, however, that our optimization algorithm using the total-variation regularization does not rely on the proposed aperture rectification and other existing leaf-sequencing algorithms can be used here as well.

The aperture rectification algorithm modifies the optimized beamlet intensity. One may be concerned that the rectification function drives the beamlet intensity away from the optimal solution, and degrades the performance of the optimization. As shown in the section of results, with a proper choice of the penalty weight β for the regularization term, the optimized beamlet intensity is already very close to piecewise constant, and the intensity modification due to the aperture rectification function is small.

2.3. Evaluation

The proposed algorithm has been tested on a prostate patient and a head and neck patient. The algorithm was implemented in Matlab, using the MOSEK optimization software package (<http://www.mosek.com>). The optimization problem was first reformulated into a standard quadratic form as shown in appendix A. The standard quadratic optimization routine provided in MOSEK was then called to solve the problem, using an interior-point optimizer. The resultant dose distributions were evaluated using clinical acceptance criteria, such as maximum and mean doses and dose–volume constraints.

For the prostate patient, five fields were used at angles of 35°, 110°, 180°, 250° and 325°, based on a standard clinical protocol. Each field targeted the center of PTV, and contained 20 by 16 beamlets, with a beamlet size of 5 mm by 5 mm at the source-to-axis distance (SAD). To save computation in the MC simulation of the dose distribution, the CT data were downsampled in the dose calculation, and the voxel size was 3.92 mm by 3.92 mm by 2.5 mm. The rectum, bladder and femoral heads were included as sensitive structures. For a better comparison, besides the proposed algorithm, two more algorithms were implemented. The first was a beamlet-based optimization. To show the best possible plan, no leaf-sequencing algorithm was applied, which resulted in a plan using a total number of 1600 segments (one segment for each beamlet). The second was an optimization using quadratic smoothing to reduce the field complexity. To demonstrate the advantage of the proposed method, in our implementation, we simply changed the absolute values in the second term of the objective function (3) to squares.

For the head and neck patient, seven fields were used at angles of 20°, 120°, 145°, 180°, 215°, 240° and 340°, and each field contained 16 by 20 beamlets. All the other parameters of the dose calculation were the same as used on the prostate patient. The sensitive structures included the brainstem, optic chiasm/nerves, optic lens, left parotid, larynx and spinal cord. The mandible and right parotid were not used because these structures significantly overlapped with the PTV.

3. Results

3.1. Prostate results

After the dose distribution of each beamlet was calculated using the MC simulation, the proposed algorithm took about 2 min on a 3 GHz PC for the prostate plan. Figure 1 shows one example of the optimized beamlet intensity of the second incident field before the aperture rectification is applied, with no regularization, with the quadratic smoothing and with the total-variation regularization. The quadratic smoothing reduces the complexity of the intensity map to some extent, and it also smoothes the sharp edges. The total-variation regularization further reduces the field complexity by preserving the edges, and the resulting intensity map is close to a piecewise constant function. This effect is better illustrated in figure 2, which plots the

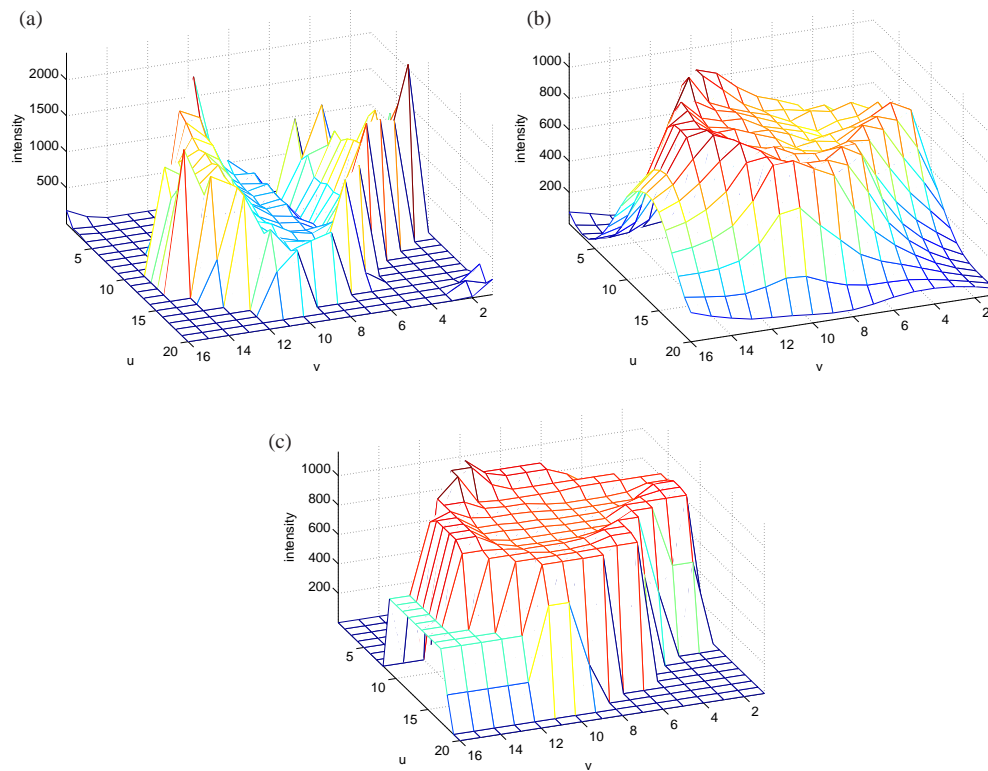


Figure 1. Optimized beam intensity maps with no regularization, with the quadratic smoothing and with the total-variation regularization. The data are from the second field of the prostate study before the aperture rectification is applied. (a) With no regularization ($\beta = 0$ in the objective function (3)). (b) With the quadratic smoothing. (c) With the total-variation regularization.

histograms of figure 1. For a better illustration, the dominant zero signals of the intensity map are not used in the histogram calculation.

Figure 3 shows the actual fluence maps obtained using the proposed optimization and aperture rectification for the prostate patient. The resultant plan using our method has different levels of modulation for different fields. In this case, the numbers of segments are 5, 6, 11, 10 and 3, for fields 1–5. Using a commercial treatment planning system (Eclipse, Varian Medical Systems, Palo Alto, CA), the numbers of segments for the fields are 11, 12, 13, 12 and 13, respectively. Our algorithm reduces the total number of segments from 61 to 35. Figure 4 shows the dose–volume histograms (DVHs) of the PTV for the prostate plan, when different total numbers of segments (N_t) are used. The plans are normalized such that 95% of the PTV volume receives 100% prescribed dose (78 Gy). As N_t increases, PTV coverage improves. The result using a beamlet-based optimization without delivery constraints is also included in the plot. Figure 5 shows the DVHs of the PTV and the sensitive structures for $N_t = 35$ using the beamlet-based method without regularization, using the quadratic smoothing and using the total-variation regularization. The proposed aperture rectification algorithm is used in the leaf-sequencing step to achieve the same number of segments in these three approaches. It is clear that the total-variation regularization achieves a better plan on the conformity of the PTV dose distribution when a small number of segments is used. Table 2 summarizes the planning results and compares with the clinical acceptance criteria. All the acceptance criteria

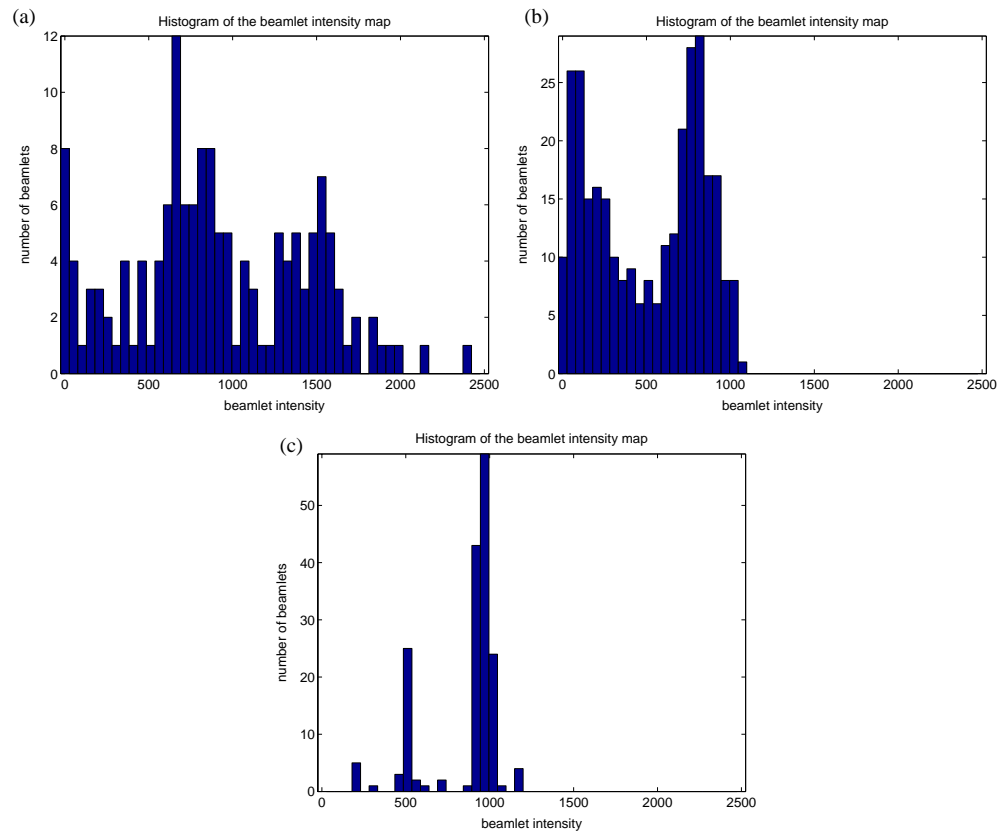


Figure 2. Histograms of the beam intensities of figure 1. Zero intensity values are not included in the calculation. (a) With no regularization ($\beta = 0$ in the objective function (3)). (b) With the quadratic smoothing. (c) With the total-variation regularization.

are satisfied. Figure 6 shows the iso-dose distributions, using the beamlet-based optimization and using the proposed algorithm. It is seen that our algorithm achieves a highly conformal dose distribution, which is comparable to the optimal result with an extremely large number of segments.

3.2. Head and neck results

The proposed optimization took about 2.5 min on a 3 GHz PC for the head and neck plan. The results of treatment planning on the head and neck patient are shown in figures 7, 8 and 9. Table 3 compares the results with the acceptance criteria and indicates that the plan is clinically acceptable. These results are achieved using numbers of segments, 1, 4, 3, 4, 4, 8 and 4 for fields 1–7. Using the eclipse system, the numbers of segments for different fields are 17, 16, 13, 15, 18, 14 and 14, respectively. Our algorithm reduces the total number of segments from 107 to 28.

It is important to let the algorithm optimize the number of segments for each field based on the field-specific modulation complexity. To illustrate this fact, figure 10 compares the DVH's of the PTV obtained using two different schemes in choosing the number of segments for each field. The first scheme uses the proposed algorithm to determine field-specific

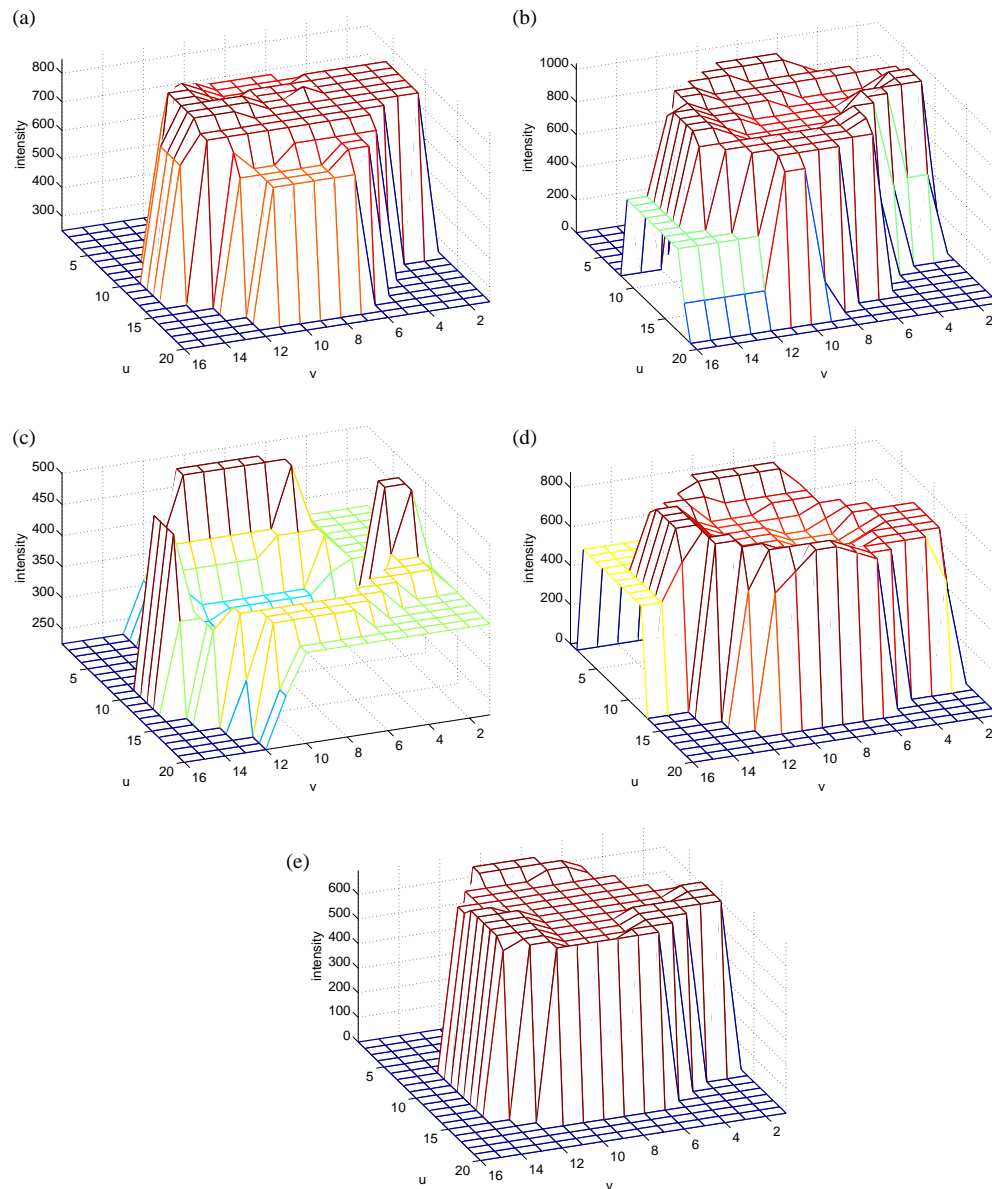


Figure 3. Actual fluence maps for the prostate plan using the proposed algorithm. (a) field 1, using five segments. (b) field 2, using six segments. (c) field 3, using 11 segments. (d) field 4, using 10 segments. (e) field 5, using three segments.

numbers of segments. The total number of segments for all fields is set to 28. The second scheme modifies the proposed algorithm. In selection of the segments, it does not consider the modulation complexity of each field and uses an equal number of four segments per field (with the same total number of segments as that in the first scheme). The plans are normalized such that 95% of the PTV volume receives 100% prescribed dose (66 Gy). It is evident that the scheme of field-specific number of segments yields a better treatment plan.

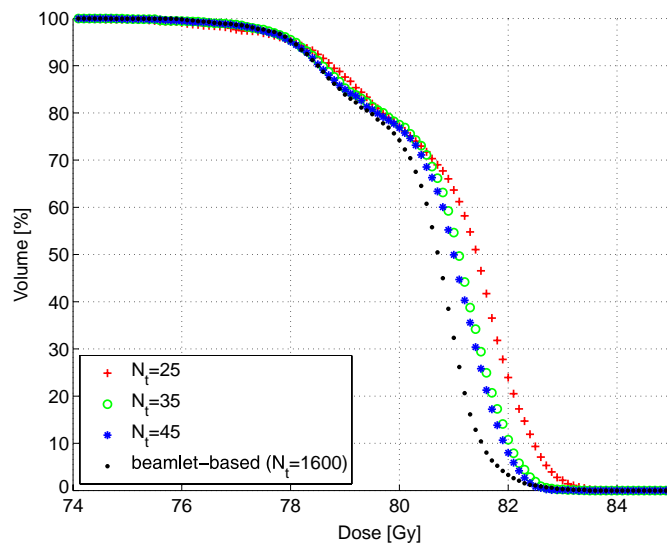


Figure 4. DVH's of the prostate PTV using the beamlet-based optimization without delivery constraints and using the proposed algorithm. Note that the total number of segments in the beamlet-based method is extremely large.

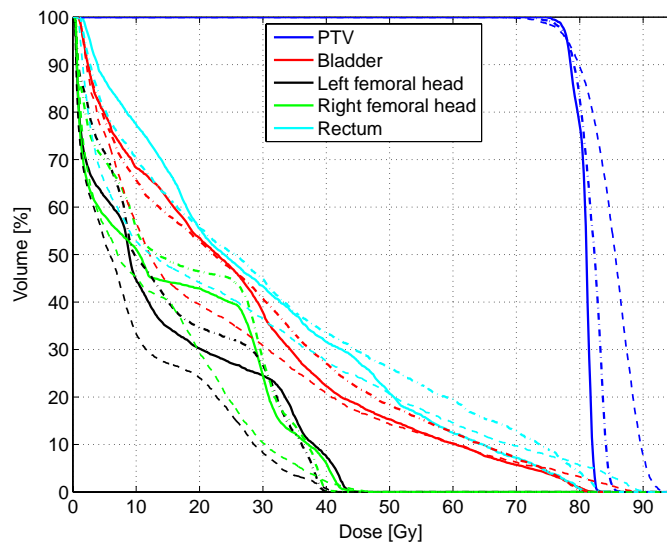


Figure 5. DVH's of the prostate plans using the beamlet-based method without regularization (dashed lines), using the quadratic smoothing (dot-dashed lines), and using the proposed algorithm (solid lines). All the plans use a total number of 35 segments ($N_t = 35$).

Figure 11 illustrates the relationship between the algorithm performance and choice of parameters, N_t and β . Note that only the first term of the objective function in the optimization (3) as in conventional planning algorithms (2) is used to quantify the algorithm performance. The result is consistent with the discussion in section 2.2.1. In general, a large N_t results in an improved dose distribution due to the increased modulation. If N_t is fixed, either a very small or a very large β value results in degraded algorithm performance. The β value that obtains

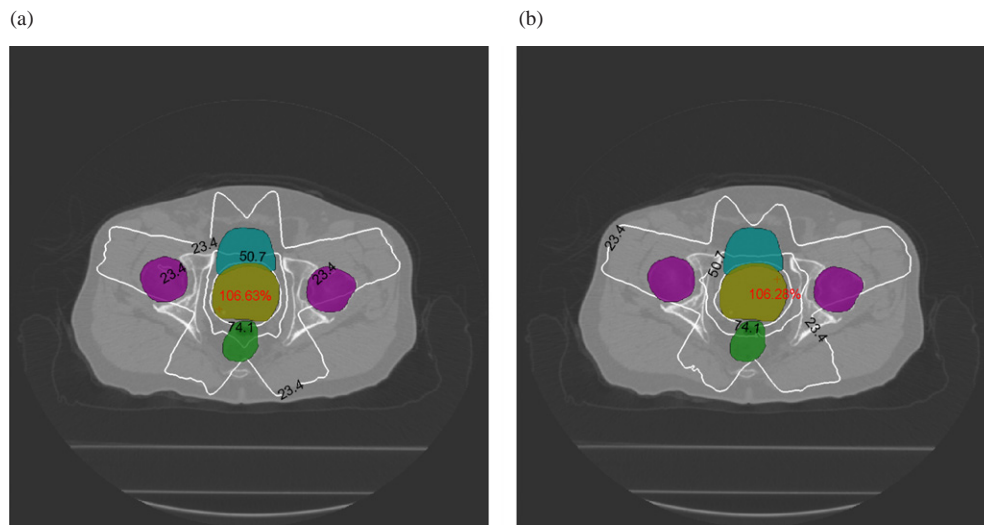


Figure 6. Dose distributions of the prostate plan. The iso-dose lines correspond to 95%, 65% and 30% of the prescribed dose (78 Gy). The PTV and the sensitive structures (bladder, rectum and femoral heads) are patched using different colors. The hotspots are marked using red crosses. (a) Using the beamlet-based optimization without delivery constraints ($N_t = 1600$). (b) Using the proposed algorithm ($N_t = 35$).

Table 2. Prostate plan objectives and results.

Regions	Acceptance criteria	Results
PTV	%vol > 78 Gy ≥ 95	%vol > 78 Gy = 95.3
Rectum	%vol > 40 Gy ≤ 35	%vol > 40 Gy = 31.6
	%vol > 65 Gy ≤ 17	%vol > 65 Gy = 9.85
	vol > 79.6 Gy $\leq 1cc$	vol > 79.6 Gy = 0.84cc
Bladder	%vol > 40 Gy ≤ 50	%vol > 40 Gy = 22.3
	%vol > 65 Gy ≤ 25	%vol > 65 Gy = 7.57
Femoral heads	%vol > 45 Gy ≤ 1	%vol > 45 Gy = 0 (left), 0.076 (right)
Body	vol > 82.7 Gy $\leq 1cc$	vol > 82.7 Gy = 0.61cc

%vol > x Gy: percentage of the volume that receives more than x Gy dose.

vol > x Gy: size of the volume that receives more than x Gy dose.

the optimal solution tends to decrease as N_t increases, which indicates that less total-variation regularization is needed as the achievable field complexity increases. Similar conclusions can be found in figure 12, which investigates the impact of choice of algorithm parameters on the final treatment plan by comparing DVHs for different N_t and β values. The plans are normalized such that 95% of the PTV volume receives 100% prescribed dose (66 Gy). In this head and neck case, for $N_t = 28$, the optimal β is around 0.8; for $N_t = 84$, the value is reduced to around 0.4.

4. Discussion and conclusions

The proposed algorithm suppresses the dispensable intensity modulation of fluence maps using total-variation regularization. The number of delivered segments is significantly

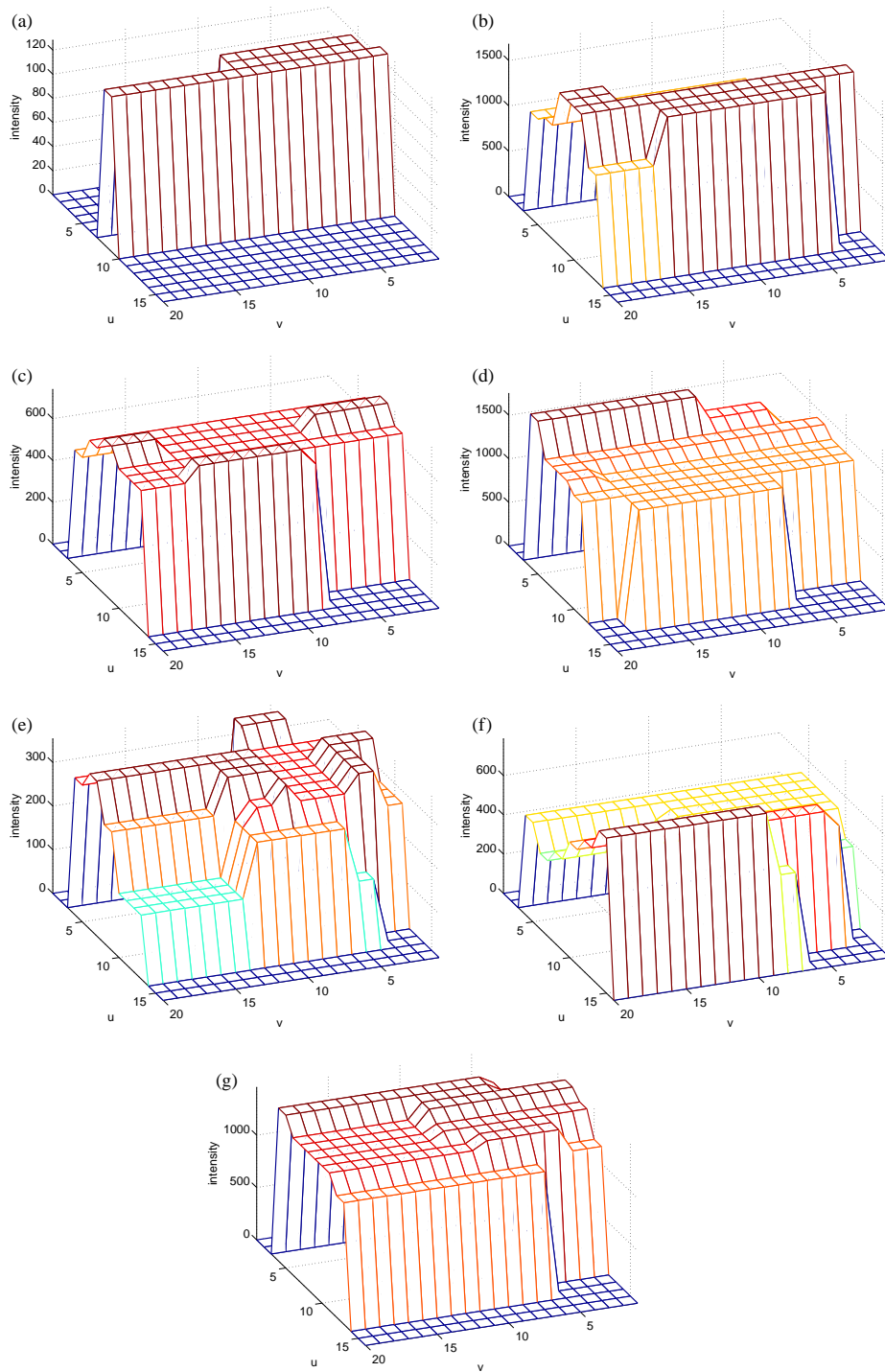


Figure 7. Actual fluence maps for the head and neck plan using the proposed algorithm. (a) Field 1, using one segment. (b) Field 2, using four segments. (c) Field 3, using three segments. (d) Field 4, using four segments. (e) Field 5, using four segments. (f) Field 6, using eight segments. (g) Field 7, using four segments.

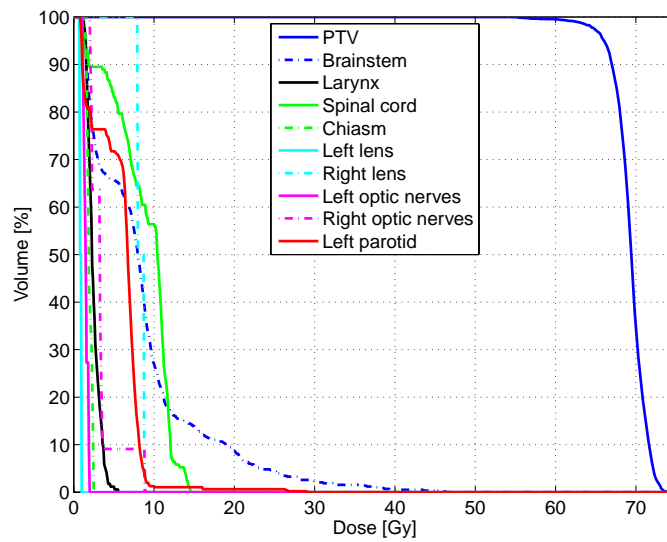


Figure 8. DVHs of the head and neck plan using the proposed algorithm ($N_t = 28$).

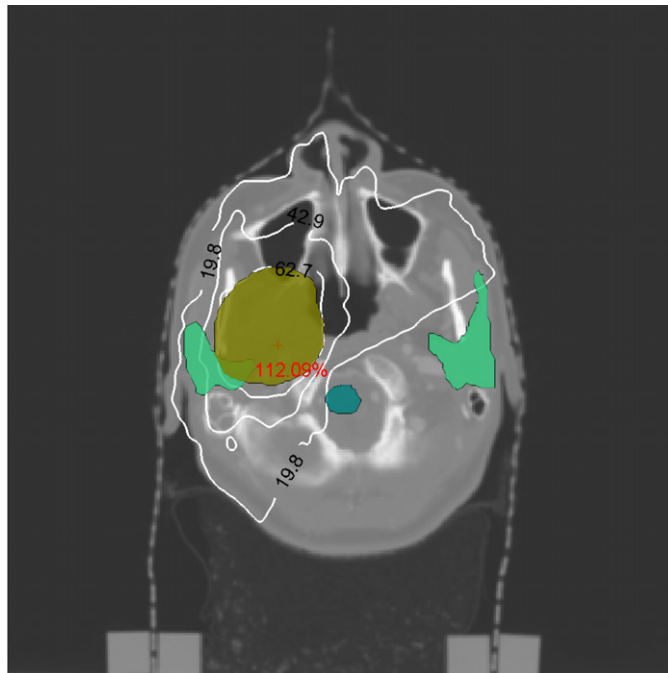


Figure 9. Dose distribution of the head and neck plan. The iso-dose lines correspond to 95%, 65% and 30% of the prescribed dose (66 Gy). The PTV and the sensitive structures (spinal cord, parotids) are patched using different colors. The hotspot is marked using a red cross.

reduced without compromising the conformity of the dose distribution. As compared to other algorithms that also include the aperture constraint in the optimization, such as DAO methods,

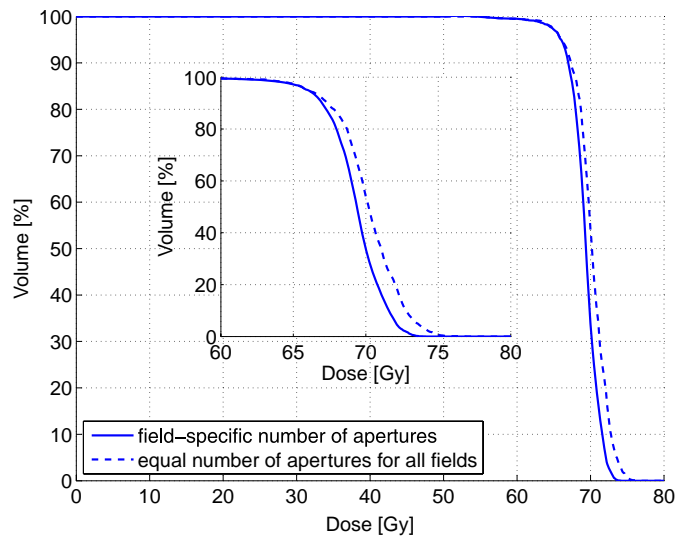


Figure 10. DVHs of the PTV in the head and neck plan using field-specific numbers of segments and using equal numbers of segments for all fields. A zoom-in plot is provided as an inset.

Table 3. Head and neck plan objectives and results.

Regions	Acceptance criteria	Results
PTV	%vol > 66 Gy \geq 95	%vol > 66 Gy = 95.0
Brainstem	Maximum < 54 Gy	Maximum = 46.4 Gy
Larynx	Maximum < 70 Gy	Maximum = 5.60 Gy
	Mean < 26 Gy	Mean = 2.42 Gy
Optic chiasm/nerves	Maximum < 54 Gy	Maximum = 8.90 Gy
Lens	Maximum < 12 Gy	Maximum = 8.80 Gy
Left parotid	Maximum < 70 Gy	Maximum = 29.1 Gy
	Mean < 26 Gy	Mean = 5.80 Gy
Spinal cord	Maximum < 45 Gy	Maximum = 14.6 Gy
Body	Maximum < 75.9 Gy (115%)	Maximum = 74.0 Gy (112.0%)

%vol > x Gy: percentage of the volume that receives more than x Gy dose.

our method has an advantage of easy implementation and high computational efficiency. The implementation of the optimization step requires a simple adaptation of the conventional algorithm, and it introduces only one additional empirical parameter β . The final form of the optimization is quadratic and efficient computation is possible using standard software packages. Furthermore, our algorithm optimizes the individual number of segments for different fields, based on the complexity of the required modulation. As shown in the result of the head and neck plan (figure 10), a treatment plan with field-specific numbers of segments greatly improves the conformity of the dose distribution. In most DAO methods, numbers of segments for each field need to be specified before the calculation, and they are not included in the optimization. Therefore, the optimality of the final solution is compromised.

A quadratic objective function is used in this study for the purpose of demonstration. The proposed total-variation regularization does not require the objective to be quadratic, and can be easily used in other forms of optimization, such as linear programming (Romeijn

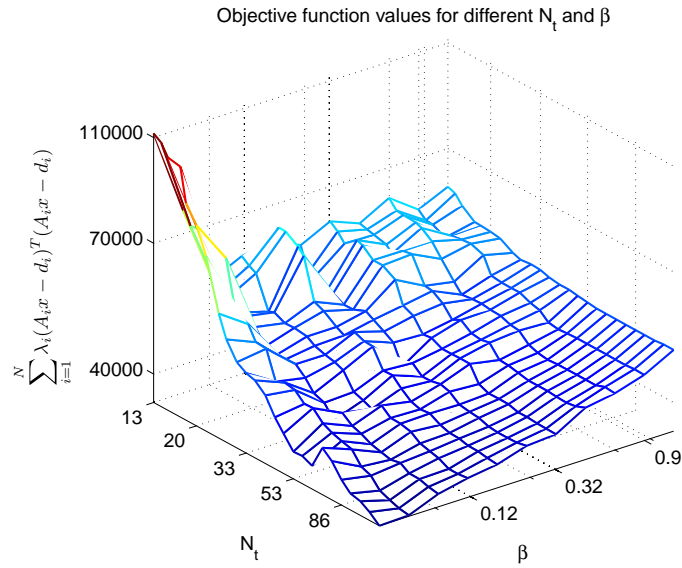


Figure 11. The value of the objective function (expression (2)) as N_t and β change. Note that all the axes are plotted in the log scale.

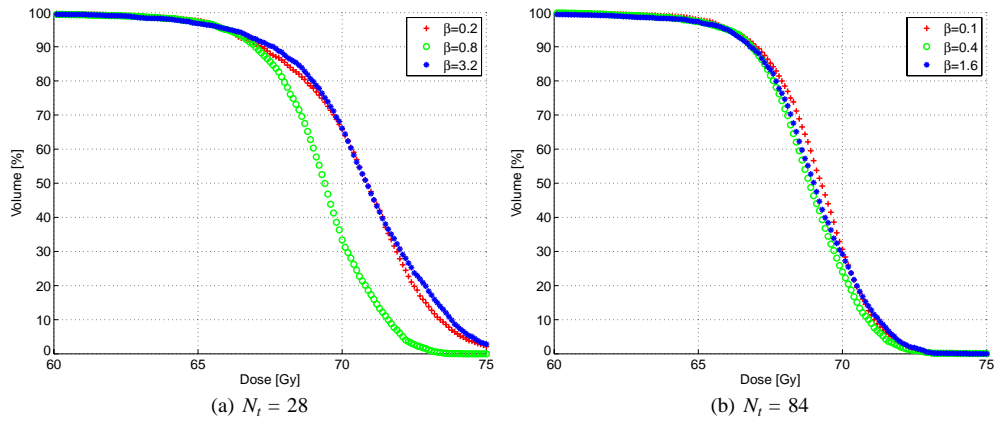


Figure 12. DVHs of the PTV in the head and neck plan using the proposed algorithm with different β and N_t values.

et al 2003). The relative weight β associated with the term of total-variation regularization is critical to the algorithm performance. In this work, we choose the value of β empirically based on general considerations of the intensity map complexity and the total number of segments. The optimal β , however, is hard to find analytically, mainly due to the non-convexity of this problem. If computation time is not an issue, the β value can be further refined adaptively using iterations, in a similar way as in the optimization of the structure specific importance factors (Xing *et al 1999*).

Our algorithm obtains a reduced number of delivered segments by reducing the complexity of the field intensity maps. The algorithm is distinct from the existing methods which use

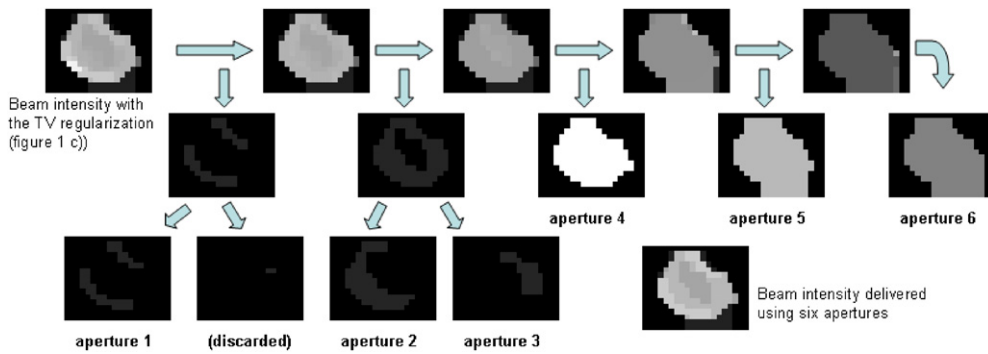


Figure 13. An example of steps 4–6 of the aperture rectification algorithm. Six apertures are derived from the beamlet intensity map shown in figure 1(c). The top row shows the residual beam intensity as aperture intensity maps are extracted. For a better illustration, different display windows are used. The middle row shows the single-level beamlet intensity maps, of which some maps are not deliverable using apertures. The same display windows are used. The first four images of the bottom row are apertures generated from undeliverable single-level intensity maps. Note the second image is discarded due to its low total intensity.

smoothing techniques, such as quadratic smoothing (Alber and Nsslin 2000, Matuszak *et al* 2007, Spirou *et al* 2001), to reduce the field complexity. Instead of smoothing the field intensity maps, we focus on shaping the intensity maps to be piecewise constant such that they can be delivered using a small number of segments. The optimized intensity maps contain sharp transitions, which are otherwise smoothed if quadratic smoothing is used. Total-variation regularization is used in our approach, because the method is able to heuristically find a sparse solution. In the segment-based treatment planning problem, the deliverable fluence maps using a small number of segments are sparse representations in the space of all fluence maps if the MLC constraints are not considered. The advantage of using total-variation regularization over using quadratic smoothing is demonstrated in the comparisons of clinical results. The difference between quadratic smoothing and total-variation regularization can also be found in classic textbooks (Boyd and Vandenberghe 2004), and they have different advantages in different applications.

In summary, a total-variation based inverse planning method is proposed in this work. The method uses quadratic optimization with the total-variation regularization, followed by an aperture rectification. As compared to other existing methods, the proposed algorithm is derived using different principles and implemented efficiently. The patient studies show that the proposed algorithm significantly reduces the total number of segments used in the treatment without compromising the delivered dose distribution.

Acknowledgments

The authors would like to thank Dr Jiji Antony for her help with the clinical data. This project is supported partially by National Cancer Institute (1R01 CA98523 and CA104205) and Department of Defense (PC046282).

Appendix A. Formulation of the optimization as quadratic programming

The first term of the optimization objective function (expression (3)), F_1 , is rewritten as

$$F_1 = \frac{1}{2}x^T Qx + c^T x + \sum_{i=1}^N \lambda_i d_i^T d_i \quad (\text{A.1})$$

where, $\sum_{i=1}^N \lambda_i d_i^T d_i$ is a constant and can be dropped in the objective function, and,

$$Q = 2 \sum_{i=1}^N \lambda_i A_i^T A_i \quad (\text{A.2})$$

$$c = -2 \sum_{i=1}^N \lambda_i A_i^T d_i. \quad (\text{A.3})$$

The second term of expression (3), F_2 , can be rewritten as

$$F_2 = \beta e^T |Bx|, \quad (\text{A.4})$$

where e is an all-one vector, with a size of $((N_u - 1)N_v N_f + N_u(N_v - 1)N_f)$ by 1, i.e., $e^T = (1, 1, 1, \dots, 1)$, $e \in \mathbb{R}^{((N_u-1)N_v N_f + N_u(N_v-1)N_f) \times 1}$; the operation of the absolute value is performed on every element of the vector Bx ; Multiplication by matrix B calculates the derivatives of x in the vertical and horizontal directions. Specifically,

$$B = \begin{bmatrix} B_h \\ B_v \end{bmatrix}, \quad (\text{A.5})$$

where B_h is used to calculate the derivatives in the u direction

$$B_h = \begin{bmatrix} S_1 & & & 0 \\ & S_2 & & \\ & & \ddots & \\ 0 & & & S_{N_v N_f} \end{bmatrix}. \quad (\text{A.6})$$

S_i are identical, with a size of $(N_u - 1)$ by N_u

$$S_i = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}. \quad (\text{A.7})$$

B_v is used to calculate the derivatives in the v direction

$$B_v = \begin{bmatrix} T_1 & & & 0 \\ & T_2 & & \\ & & \ddots & \\ 0 & & & T_{N_f} \end{bmatrix}. \quad (\text{A.8})$$

T_i are identical, with a size of $N_u(N_v - 1)$ -by- $N_u N_v$

$$T_i = \begin{bmatrix} -1 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{bmatrix}. \quad (\text{A.9})$$

On each row, -1 and 1 are separated by $N_u - 1$ zeros.

Define a new variable t as

$$t = |Bx|. \quad (\text{A.10})$$

According to equations (A.1) and (A.4), the optimization problem can be reformulated as

minimize

$$\frac{1}{2}x^T Qx + c^T x + \beta e^T t \quad (\text{A.11})$$

subject to

$$\begin{aligned} x &\geq 0 \\ |Bx| &= t \end{aligned}$$

which is equivalent to

minimize

$$\frac{1}{2}x^T Qx + c^T x + \beta e^T t \quad (\text{A.12})$$

subject to

$$\begin{aligned} x &\geq 0 \\ Bx - t &\leq 0 \\ Bx + t &\geq 0. \end{aligned}$$

The optimization problem is now in a form of quadratic programming. Note that for different optimization software packages, the problem can be reformulated differently for a better computation performance.

Appendix B. Aperture rectification from the intensity map

To enforce the aperture constraint on the optimized beamlet intensity maps, we first divide each of the optimized beamlet intensity maps into single-level intensity maps. These maps are then divided into deliverable apertures. Specifically, the aperture rectification algorithm is described in the following steps:

Step 1. Calculate the histograms of the beamlet intensity maps (x) obtained from the optimization with the total-variation regularization (see figure 2(c)).

Step 2. Take out the histogram bins corresponding to zero beamlet intensity values, and then normalize the rest of bins of each field by the maximum bin value of that field.

Step 3. Sort the histogram bins for all fields in descending order, and select the first N_i bins.

Step 4. For each field, sort the selected n intensity values for that field in descending order, denote as I_1, I_2, \dots, I_n . Initialize a 2D beamlet intensity map y as the optimized beamlet intensity map for that field obtained from the optimization with the total-variation regularization. For the index i starting from 1:

Step 4.1. Generate a 2D map m with ones at pixels of y that are larger than $\frac{I_i + I_{i+1}}{2}$ (or, $\frac{I_n}{2}$ if $i = n$), and zeros elsewhere.

Step 4.2. The i th single-level intensity map of the field is $(I_i - I_{i+1})m$ (or, $I_n m$ if $i = n$).

Step 4.3. If $i < n$, update y by assigning I_{i+1} to pixels where m have nonzero values, update i as $i = i + 1$.

Step 5. Check whether the obtained single-level intensity maps from step 4 are deliverable using one aperture. If not, separate the single-level intensity map into multiple apertures.

Step 6. Due to the separation in step 5, we have more than N_t apertures. These apertures are then sorted according to their total intensities, and the largest N_t apertures are selected as the final deliverable apertures.

In the above description, we assume the total number of apertures N_t is pre-determined. Note that the numbers of apertures for different fields can be different and they are optimized based on the modulation complexity. We also ignore MLC transmission and MLC header scatter in this study. As in the conventional leaf-sequencing algorithms, these effects can be readily included (Xing and Li 2000, Yang and Xing 2003).

Figure 13 shows an implementation example of the aperture rectification algorithm using the second field intensity map of the prostate study (figure 1(c)). The aperture rectification algorithm first generates single-level beamlet intensity maps for all fields, based on the histogram data as shown in figure 2(c). These single-level beamlet intensity maps are then divided into deliverable apertures.

References

- Alber M and Nsslin F 2000 Intensity modulated photon beams subject to a minimal surface smoothing constraint *Phys. Med. Biol.* **45** N49–52
- Bedford J L and Webb S 2007 Direct-aperture optimization applied to selection of beam orientations in intensity-modulated radiation therapy *Phys. Med. Biol.* **52** 479–98
- Bergman A M, Bush K, Milete M P, Popescu I A, Otto K and Duzenli C 2006 Direct aperture optimization for IMRT using Monte Carlo generated beamlets *Med. Phys.* **33** 3666–79
- Block K T, Uecker M and Frahm J 2007 Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint *Magn. Reson. Med.* **57** 1086–98
- Bortfeld T 1999 Optimized planning using physical objectives and constraints *Semin. Radiat. Oncol.* **9** 20–34
- Bortfeld T R, Kahler D L, Waldron T J and Boyer A L 1994 X-ray field compensation with multileaf collimators *Int. J. Radiat. Oncol. Biol. Phys.* **28** 723–30
- Boyd S and Vandenberghe L 2004 *Convex Optimization* (Cambridge: Cambridge University Press) pp 310–7
- Breedveld S, Storchi P R M, Keijzer M and Heijmen B J M 2006 Fast, multiple optimizations of quadratic dose objective functions in IMRT *Phys. Med. Biol.* **51** 3569–79
- Cho P S and Phillips M H 2001 Reduction of computational dimensionality in inverse radiotherapy planning using sparse matrix operations *Phys. Med. Biol.* **46** N117–25
- Cotrutz C and Xing L 2003 Segment-based dose optimization using a genetic algorithm *Phys. Med. Biol.* **48** 2987–98
- Ezzell G A, Galvin J M, Low D, Palta J R, Rosen I, Sharpe M B, Xia P, Xiao Y, Xing L, Yu C X, (subcommittee IMRT and committee AAPMRT) 2003 Guidance document on delivery, treatment planning, and clinical implementation of IMRT: report of the IMRT Subcommittee of the AAPM Radiation Therapy Committee *Med. Phys.* **30** 2089–115
- Kamath S, Sahni S, Palta J, Ranka S and Li J 2004 Optimal leaf sequencing with elimination of tongue-and-groove underdosage *Phys. Med. Biol.* **49** N7–19
- Kawrakow I 1997 Improved modeling of multiple scattering in the voxel Monte Carlo model *Med. Phys.* **24** 505–17
- Kawrakow I, Fippel M and Friedrich K 1996 3D electron dose calculation using a voxel based Monte Carlo algorithm (VMC) *Med. Phys.* **23** 445–57
- Kolehmainen V, Vanne A, Siltanen S, Järvenpää S, Kaipio J P, Lassas M and Kalke M 2006 Parallelized Bayesian inversion for three-dimensional dental x-ray imaging *IEEE Trans. Med. Imaging* **25** 218–28
- Kuterdem H G and Cho P S 2001 Leaf sequencing with secondary beam blocking under leaf positioning constraints for continuously modulated radiotherapy beams *Med. Phys.* **28** 894–902

- Ma L 2002 Smoothing intensity-modulated treatment delivery under hardware constraints *Med. Phys.* **29** 2937–45
- Ma L, Boyer A L, Ma C M and Xing L 1999 Synchronizing dynamic multileaf collimators for producing two-dimensional intensity-modulated fields with minimum beam delivery time *Int. J. Radiat. Oncol. Biol. Phys.* **44** 1147–54
- Matuszak M M, Larsen E W and Fraass B A 2007 Reduction of IMRT beam complexity through the use of beam modulation penalties in the objective function *Med. Phys.* **34** 507–20
- Mestrovic A, Milete M P, Nichol A, Clark B G and Otto K 2007 Direct aperture optimization for online adaptive radiation therapy *Med. Phys.* **34** 1631–46
- Michalski D, Xiao Y, Censor Y and Galvin J M 2004 The dose–volume constraint satisfaction problem for inverse treatment planning with field segments *Phys. Med. Biol.* **49** 601–16
- Oelfke U and Bortfeld T 1999 Inverse planning for x-ray rotation therapy: a general solution of the inverse problem *Phys. Med. Biol.* **44** 1089–104
- Romeijn H E, Ahuja R K, Dempsey J F and Kumar A 2005 A column generation approach to radiation therapy treatment planning using aperture modulation *SIAM J. Opt.* **15** 838–62
- Romeijn H E, Ahuja R K, Dempsey J F, Kumar A and Li J G 2003 A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planning *Phys. Med. Biol.* **48** 3521–42
- Rudin L I, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms *Phys. D* **60** 259–68
- Saw C B, Siuchi R C, Ayyangar K M, Zhen W and Enke C A 2001 Leaf sequencing techniques for MLC-based IMRT *Med. Dosim.* **26** 199–204
- Shepard D M, Earl M A, Li X A, Naqvi S and Yu C 2002 Direct aperture optimization: a turnkey solution for step-and-shoot IMRT *Med. Phys.* **29** 1007–18
- Spirou S V, Fournier-Bidoz N, Yang J, Chui C S and Ling C C 2001 Smoothing intensity-modulated beam profiles to improve the efficiency of delivery *Med. Phys.* **28** 2105–12
- Sun X and Xia P 2004 A new smoothing procedure to reduce delivery segments for static MLC-based IMRT planning *Med. Phys.* **31** 1158–65
- Thieke C, Nill S, Oelfke U and Bortfeld T 2002 Acceleration of intensity-modulated radiotherapy dose calculation by importance sampling of the calculation matrices *Med. Phys.* **29** 676–81
- van Asselen B, Schwarz M, van Vliet-Vroegindeweij C, Lebesque J V, Mijnheer B J and Damen E M F 2006 Intensity-modulated radiotherapy of breast cancer using direct aperture optimization *Radiother. Oncol.* **79** 162–9
- Webb S, Convery D J and Evans P M 1998 Inverse planning with constraints to generate smoothed intensity-modulated beams *Phys. Med. Biol.* **43** 2785–94
- Xia P and Verhey L J 1998 Multileaf collimator leaf sequencing algorithm for intensity modulated beams with multiple static segments *Med. Phys.* **25** 1424–34
- Xiao Y, Michalski D, Censor Y and Galvin J M 2004 Inherent smoothness of intensity patterns for intensity modulated radiation therapy generated by simultaneous projection algorithms *Phys. Med. Biol.* **49** 3227–45
- Xing L and Li J G 2000 Computer verification of fluence map for intensity modulated radiation therapy *Med. Phys.* **27** 2084–92
- Xing L, Li J G, Donaldson S, Le Q T and Boyer A L 1999 Optimization of importance factors in inverse planning *Phys. Med. Biol.* **44** 2525–36
- Xing L, Wu Y, Yang Y and Boyer A 2005 Physics of intensity modulated radiation therapy *Intensity Modulated Radiation Therapy* ed A J Mundt and J C Roeske (Hamilton: Decker Inc.) pp 20–52
- Yang Y and Xing L 2003 Incorporating leaf transmission and head scatter corrections into step-and-shoot leaf sequences for IMRT *Int. J. Radiat. Oncol. Biol. Phys.* **55** 1121–34

IMRT Optimization through Voxel-based Prescription Correction

Jordan LeNoach, Lei Zhu, Pavel Lougovski, Yunzhi Ma, Lei Xing

Department of Radiation Oncology, Stanford University, Stanford, CA 94305 USA

Abstract: The success of IMRT inverse treatment plans depends greatly upon how well the dosimetric goals of different structures within a patient are balanced and prioritized. In most plans, the optimal beam intensity profiles are calculated by assigning a uniform importance and prescription to all voxels in the target volume and sensitive structures at hand. While a perfectly uniform and accurate dose is desired in all structures, this is impossible to achieve in reality and many voxels will receive a different dose than their given prescriptions. Different methods have been explored by which importance factors can be modified or removed from optimization to improve treatment plans. This paper explores an alternative of assigning individual voxel prescriptions and modifying them between optimizations in order to produce an improved dose distribution. A clinical head and neck case was used to test this method. By correcting voxel prescriptions to compensate for the inequalities between the actual and desired doses calculated, substantial improvements are obtained for the treatment plan as large dose reductions were achieved in almost all of the critical structures present.

Working Manuscript – in progress.

Introduction: IMRT treatment planners will typically formulate radiation delivery into a linear or quadratic optimization problem in order to achieve fluence maps of beamlet intensities that will deliver an acceptable dose to all structures. The problem is complicated by the unequal importance of different goals; the achievement of the desired dose in the tumor region is often of much greater importance than reducing the dose received by critical organs, and different organs vary in their dose toleration. The solution in most plans is to multiply the dose distributions for each structure by an importance factor, a constant that can either increase or decrease the penalty for error between the desired and delivered dose. Through the modification of importance factors, one can achieve a treatment plan much more compliant with the given dosimetric goals of a patient. However, since patient cases vary greatly, there is not one set of importance factors that will produce an optimal treatment plan for everyone. Planners must often go through a tedious and time-consuming process of repeatedly changing importance factors and re-optimizing beamlet intensities until an acceptable distribution is achieved. This current method leaves much to be desired. One of the largest problems is that it considers dose tradeoffs at the structural level only. Yet, the voxels within the target volume and critical structures are not equal and can oppose each other in terms of their dosimetric goals. As a result, assigning them all equal importance and prescriptions can greatly limit the solution space of the optimization. There are a number of proposed methods that can be used to take into account this inequality among voxels. For instance, as explored by Yang and Xing, one can implement voxel-specific importance factors rather than using a single, constant importance factor for a structure as justified by dosimetric capacities [1,3]. Others have explored methods by which to eliminate the need for importance factors altogether. In this paper, an alternative is explored by which voxels are assigned individual prescriptions as opposed to a single prescription for an entire structure. These

prescriptions are iteratively modified prior to optimization to produce a new solution. Each voxel is given a higher or lower prescription based on how well it achieved the desired dose in the previous optimization. This method of prescription correction offers advantages over traditional IMRT planning in that it forces the solution closer to the optimal dose with greater uniformity in a very intuitive manner. For the purpose of demonstration, this approach was implemented with the method proposed by Zhu et al. for total-variation regularization [2]. Due to the simplicity of the approach, it can easily be added to almost any other treatment planning algorithm.

Method: We formulate our method as an iterative procedure. Initially, all prescriptions are set to their ideal value for both organs and target volume. Next, the optimal beamlet intensities for the given parameters are calculated as a quadratic optimization problem. Once the delivered doses are determined, the prescriptions are corrected linearly based on deviation from the delivered dose with the ideal prescription. Finally, the beamlet intensities are re-optimized with the new prescriptions. This process can be repeated until the solution converges.

For the patient case used, the objective function was reformulated and solved as a quadratic optimization problem with the MOSEK optimization software package. The dose distribution produced by a series of beamlet intensities is given by the relation

$$d = Ax$$

where d is the delivered dose distribution, x is a series of beamlet intensities, and A is a matrix that relates beamlet intensities to voxel doses within a given structure. The A matrix is calculated based on the CT image of a patient, the beam geometry, and the treatment machine. The objective function to be minimized is given by

$$\sum_{\sigma=1}^N r_{\sigma} F_{\sigma}$$

Where r represents an importance factor and the objective function for each structure is given by

$$E_s = \sum_{i=1}^{N_s} [D_c(i) - D_0(i)]^2$$

D_c represents the calculated dose and D_0 is the prescribed dose for a voxel i . The objective function calculates the square of the error between the prescribed and received doses of all voxels in a given structure. The prescription corrections are then performed after each optimization. A uniform target dose is assigned for each structure and remains unchanged for all iterations, but the actual prescriptions for each voxel are assigned a higher or lower dose for the next optimization iteration depending on the error between the delivered dose and the target dose. A linear function was chosen since higher order functions are likely to overcorrect and produce unstable solutions. Each voxel is corrected by the following equation

$$D_{n+1}(i) = \gamma_s (T_s - D_c(i)) +$$

where i indicates a single voxel in structure s , D_0 is the prescribed dose for iteration n , T is the target dose for a given structure, and D_c is the delivered dose calculated from the last optimization. The constant γ_s determines what proportion of the error will be used to correct the prescription. Structures with a higher value will better achieve their target dose with a quicker rate of correction, usually at the expense of others with lower values. However, setting too large a value may result in unstable solutions, which is why it best kept between 0 and 1. To implement total-variation regularization, a second term was added to the objective function that

calculates the horizontal and vertical derivatives of the beamlet intensities for each field. With this term, the beamlet intensities are forced into piece-wise constant fluence maps with a small number of segments, making the fluence maps easier to deliver.

A clinical head and neck case was used to test this method of prescription corrections. For the planning target volume (PTV), the target dose was set to 66 Gy. The beamlet intensities are normalized so that 95% of PTV receives 66 Gy for every dose distribution. All organs at risk (OAR) including the tuning structure were assigned target doses of zero. In this study, the proportionality constants were set to one for all structures with the exception of the tuning structure. Since the tuning structure surrounds the PTV, it was given a smaller constant so that its prescription corrections would not negate those of the PTV, yet still prevent greater irradiation to surrounding tissue. However, it should be high enough to prevent severe irradiation to surrounding tissue.

Results: The optimization algorithm described was implemented and tested on a head and neck case. The dose volume histogram (DVH) for the treatment plan in figure 1 shows substantial improvement. While prescription correction resulted in a minor improvement to the PTV, most of the critical structures received a dramatic reduction in dose. This can be attributed to the tradeoff determined by the importance factors assigned to each structure during optimization. Because the PTV's importance factor is much higher relative to every other structure, it will achieve a nearly optimal dose without prescription corrections. However, the OARs can be much further from their optimal dose and repeated prescription corrections allow the OARs to be forced to a lower dose. Figure 2 shows the isodose distribution with contour lines at 95%, 65%, and 30% of the PTV target dose (66 Gy). The distribution shows that there is a great reduction of radiation to the surrounding tissue as well. The right parotid and mandible were not included

in the DVH of the head and neck case because they overlapped with the PTV. Table 1 shows the results based on the criteria used for determining whether a treatment plan is clinically acceptable. The PTV and OARs in the corrected case all meet the standard requirements for head and neck treatment plans. Due to the large improvements, this method of prescription correction can ease the process producing a clinically acceptable treatment plan and may eliminate or reduce the need to manually change the importance factors on structures.

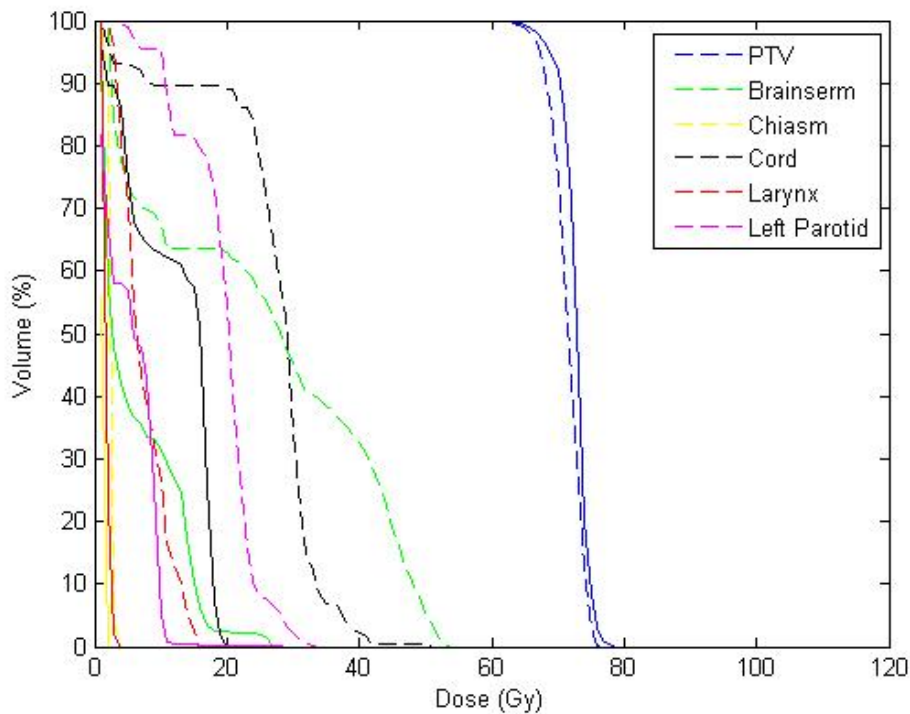


Figure 1: DVH of the head and neck plan. Dashed lines indicate the DVH before prescription correction and solid lines show it after.

Table 1: Head and Neck plan results before and after prescription corrections

Regions	Original	Corrected
PTV	% vol > 66Gy = 97.92	% vol > 66Gy = 98.66
Brainstem	maximum = 53.69Gy	maximum = 27.19Gy
Larynx	maximum = 15.94Gy	maximum = 3.26Gy
	mean = 7.39Gy	mean = 1.73Gy
Chiasm	maximum = 3.03Gy	maximum = 1.29Gy

Cord	maximum = 50.23Gy	maximum = 19.24Gy
Left Parotid	maximum = 33.49Gy	maximum = 28.99Gy
	mean = 19.2Gy	mean = 5.51Gy
Left Optic nerve	maximum = 4.4Gy	maximum = 0.93Gy
Right optic nerve	maximum = 5.25Gy	maximum = 1.54Gy
Left lens	maximum = 3.34Gy	maximum = 0.47Gy
Right lens	maximum = 3.99Gy	maximum = 0.83Gy

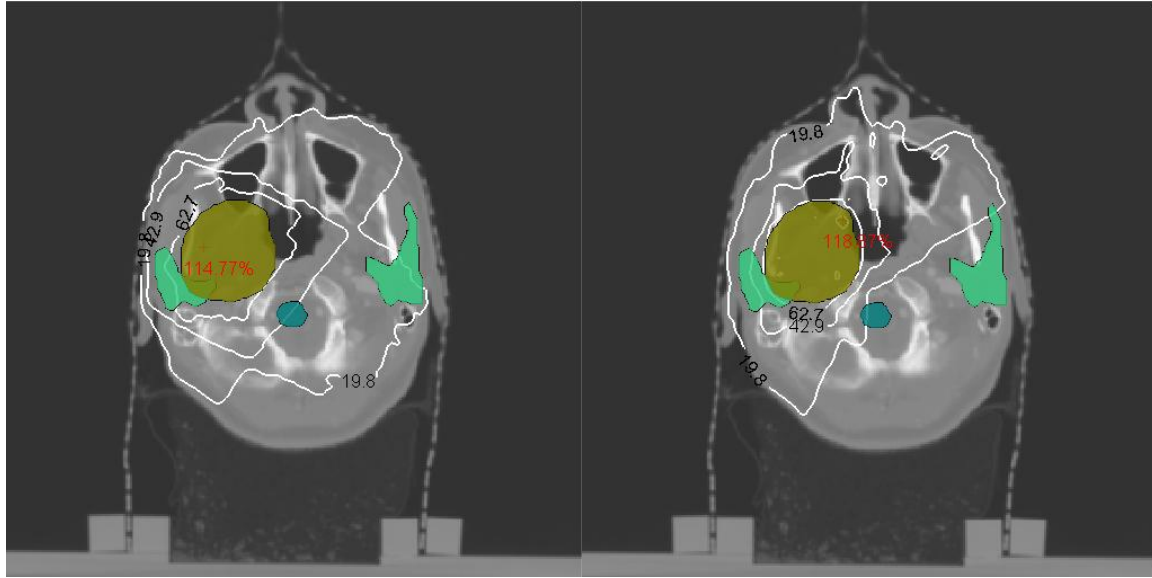
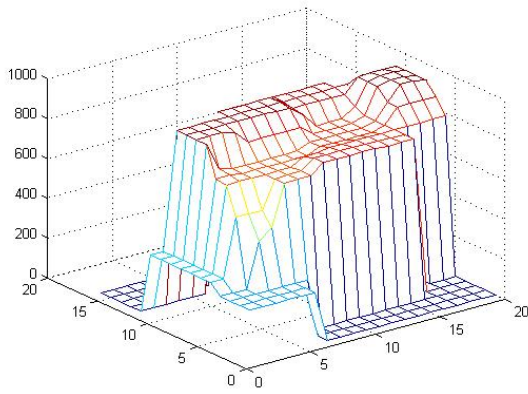


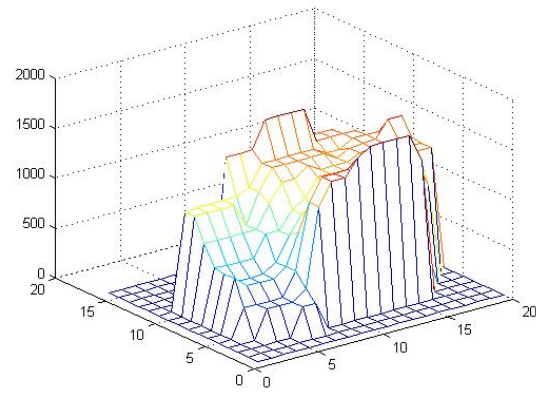
Figure 2: Head and neck plan dose distributions before prescription correction (left) and after (right). Hotspots are marked with red crosses.

In our specific implementation of prescription corrections, one of the difficulties encountered with the corrections is that they would act against the total-variation regularization term in the objective function. Because the penalties for voxels continue increasing after each iteration in order to force them to their target doses, the penalties eventually become larger than that of the regularization term, forcing them out of their piece-wise constant form and making the fluence maps more complex. Thus, after several iterations of corrections, the calculated fluence maps may require many more apertures for delivery than initially. In order to overcome this problem, the penalty weight for the term of total variation in the objective function must be set to

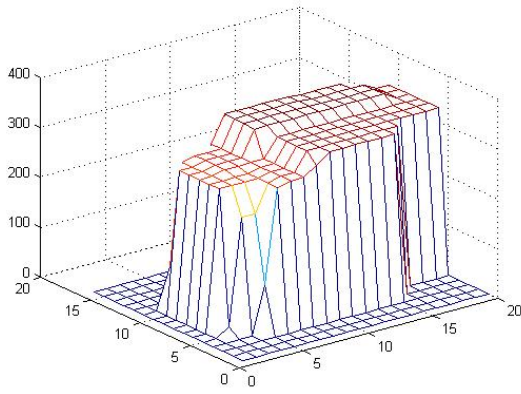
a value high enough so that the resulting fluence maps after a given number of iterations will still be deliverable with an acceptable number of apertures. For this case, it was set to 50. Figure 3 illustrates these results and shows how the prescription corrections affect the fluence map when the penalty weight is not large enough. Alternatively, one may choose to set the penalty weight for regularization to zero until the last iteration of



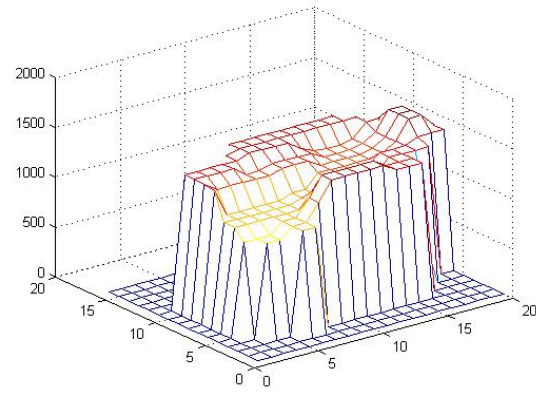
0 prescription corrections, $\beta=20$



4 prescription corrections, $\beta=20$



0 prescription corrections, $\beta=50$



4 prescription corrections, $\beta=50$

Figure 3: The fluence maps of one field before and after prescription corrections with different penalty weights for total-variation regularization.

optimization. The advantage of doing this is that it reduces computation time since the objective function is much simpler without regularization. However, this tends to result in a slightly lower quality DVH as opposed to setting a high, constant penalty weight for all iterations. It should be noted that for other treatment plans, some experimentation may be required to determine what value the penalty weight should be assigned for a given number of iterations.

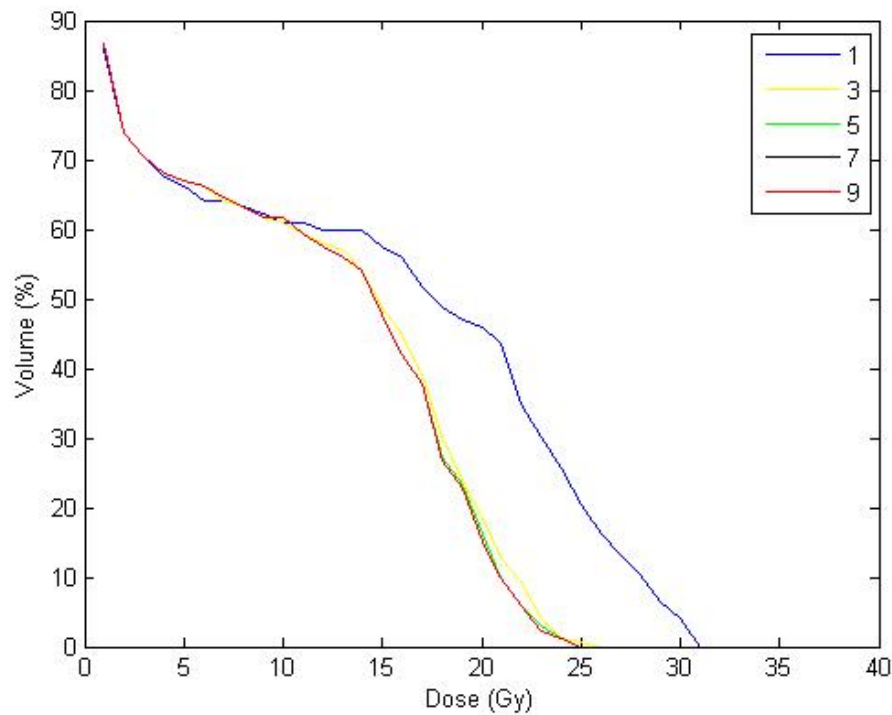


Figure 4: DVH of the spinal cord in the head and neck plan for every odd optimization iteration.

These results were obtained after five iterations of optimization (four prescription corrections) with total-variation regularization. The running time only depends on the optimization algorithm used and however many iterations of prescription corrections one wishes to execute. While more iterations typically produce better results, the prescription correction can result in significant improvements even after one or two iterations only. However, the solutions do begin to converge after a certain number of iterations as shown in figure 4. It has been

observed that the solution will become unstable after a large number of iterations (for this case, about 20), but this should not be a problem since there is no need to perform so many. The great improvement prescription corrections produced for the critical organs can most likely be attributed to the assignment of negative prescription values. Although no structure can receive a negative dose in reality, a negative prescription enables a much greater penalty for any radiation OARs are exposed to. While a regularization based optimization algorithm was used for this study, the same method of prescription correction can be easily implemented with any other algorithm.

Conclusion: Voxel-based prescription corrections produced significant improvements in the treatment plans of the cases studied. Combining this with total-variation regularization, the fluence maps achieved can be delivered efficiently with the addition of an aperture rectification algorithm while maintaining an acceptable dose distribution. Through this simple method of prescription corrections, the desired doses were better achieved in the PTV and most of the organs at risk. Due to the simplicity and intuitiveness of this approach, it offers the advantage of being easily implemented in most IMRT planning algorithms with only a few lines of code. Various improvements could be made to this method. For instance, decaying proportionality constants may perhaps be employed in order to give a large, quick correction initially while preventing the corrections from becoming unstable.

Acknowledgement: This project is supported partially by Department of Defense (PC046282) and National Cancer Institute (CA104205).

References:

- [1] Y. Yang and L. Xing, "Inverse treatment planning with adaptively evolving voxel-dependent penalty scheme," *Med. Phys.* **31**, 2839-2844 (2004).
- [2] L. Zhu, L. Lee, Y. Ma, Y. Ye, R. Mazzeo, and L. Xing, "Using total-variation regularization for segment-based dose optimization with field specific numbers of segments," 1-23
- [3] Z. Shou, Y. Yang, C. Cotrutz, D. Levy, and Lei Xing, "Quantitation of the *a priori* dosimetric capabilities of spatial points in inverse planning and its significant implication in defining IMRT solution space," *Phys. Med.* **50**, 1469-1482 (2005).

Optimizing the number of segments in radiation therapy using compressed sensing

Lei Zhu and Lei Xing

Department of Radiation Oncology, Stanford University, Stanford, CA 94305 USA.

Abstract

In current radiation therapy using Linac machines, dose is delivered using apertures. A small number of apertures (or beam segments) is desirable for high treatment efficiency as well as accuracy due to the reduced patient motion. The traditional beamlet-based treatment planning does not include the non-convex physical constraints of apertures in the optimization, and results in a large number of beam segments. This number is reduced in segment-based planning algorithms, such as direct aperture optimization (DAO) methods. However, most of the DAO methods pre-fix the segment number for each field and little research has been done in the literature to minimize the segment numbers without much compromising the dose distribution. Since the actual fluence map is a summation of piece-wise constant beam segments and it is sparse after taking derivatives, we formulate the planning as a multi-objective optimization problem and minimize the segment numbers by using compressed sensing to find a sparse solution. In addition to the objective for dose conformity on the planning target volume and avoidance of critical structures, a total variation term is included in the optimization as a second objective function. A Pareto frontier is then calculated, and the achieved dose distributions associated with the Pareto efficient points are judged using clinical acceptance criteria. The clinically acceptable dose distribution with the smallest number of segments is chosen as the final solution. The evaluation result using the proposed method on a prostate patient shows that the number of segments is greatly reduced while a satisfactory dose distribution is still achieved.

Keywords: radiation therapy, inverse planning, compressed sensing

1 Introduction

The step-and-shoot intensity-modulated radiation therapy (IMRT) uses a series of beam apertures (or segments) shaped by the multi-leaf collimator (MLC) to deliver patient dose. Currently, there are two major categories of inverse planning algorithms to optimize the dose distribution such that it closely matches the

prescribed dose. The traditional beamlet-based algorithms divide the beam’s eye view (BEV) of the planning target volume (PTV) into small beamlets [1, 2], and a convex optimization is used to optimize the beamlet intensities without considering the non-convex physical constraints of beam apertures. For a fast calculation, a linear or quadratic function is typically used as the objective. A leaf sequencing algorithm is then applied on the resultant beamlet intensity map to generate a set of deliverable beam segments such that the summation of these beam segments is approximately equal to the optimized beamlet intensity obtained from the beamlet-based optimization [3–8]. Since the constraints of beam apertures are not included in the optimization step, these algorithms usually obtain a large number of segments due to the high complexity of optimized beamlet intensity maps. To ameliorate this problem, many algorithms have been proposed in the literature using smoothing techniques to suppress the complexity of the intensity maps [9–14]. Typical examples use an additional term of sum of derivative squares [9, 11, 15], which are often referred to as quadratic smoothing or regularization in the theory of convex optimization. Since these algorithms smooth the sharp edges of the intensity field as well, the optimized beam intensity is not piece-wise constant and still cannot be delivered using a very small number of segments. Segment-based optimization algorithms, such as direct aperture optimization (DAO), include naturally the physical constraints of beam apertures in the optimization [16–22]. The leaf-sequencing step is therefore unnecessary and the achieved number of segments for a satisfactory dose distribution is less than that using a beamlet-based optimization. The downside of including non-convex constraints in the optimization, however, is the increased difficulty of finding a global optimal solution. As a result, random search algorithms, such as simulated annealing, are commonly employed. The computation is therefore intensive and requires tuning of multiple algorithm parameters in the searching and cooling schedules. The issue is exacerbated in advanced applications such as 4D and adaptive therapy treatment planning, where the number of decision variables in the optimization is greatly increased. Another disadvantage of most of the DAO methods is that the number of segments for each field is not optimized as a decision variable. Instead, it is determined before the calculation, which compromises the optimality of the final solution.

In this work, we tackle the problem of efficiently optimizing the number of beam segments in IMRT, on which little research has been done in the literature. The derivation of the proposed method is based on the fact that a beamlet intensity map which can be delivered using a small number of segments must be piece-wise constant and its derivative is sparse. A compressed sensing algorithm is used to enforce the sparsity of the optimized solution [23, 24], such that the number of beam segments is minimized. To ensure the achieved dose distribution is not compromised, we add a total-variation term into the traditional beamlet-based optimization as a second objective function and the problem is solved as a multi-objective optimization. The Pareto efficient points are calculated, among which the clinically acceptable solution with the smallest number of beam segments is selected as the final solution. The proposed method is evaluated

using a prostate patient study.

2 Method

Table 1: Variable glossary

$A(A_i)$	matrix that relates the beamlet intensity to delivered dose;
$d(d_i)$	delivered dose;
N	total number of beamlets, $N = N_u N_v N_f$;
N_f	number of fields;
N_t	TOTAL number of segments;
N_u	number of MLC leaf positions for each leaf;
N_v	number of MLC leaves per field;
x	beamlet intensity, the decision variable in the optimization;
λ_i	importance factor associated with the i'th structure;

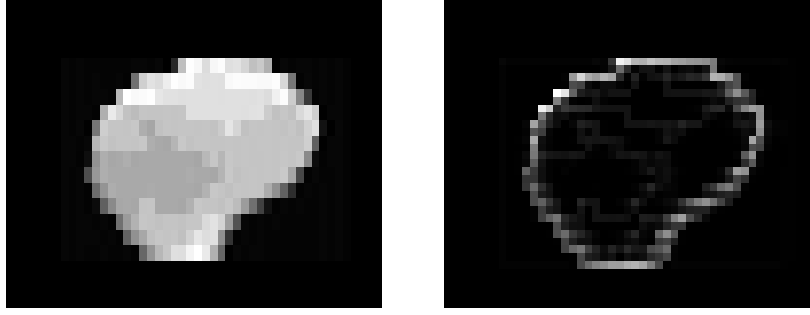
2.1 Beamlet-based optimization

The conventional beamlet-based optimization for inverse treatment planning is based on the linear relationship between the delivered dose distribution on the patient, d , and the intensity of the beamlets, x :

$$d = Ax \tag{1}$$

where d is a vectorized dose distribution for a three-dimensional volume, and the beamlet intensity x is a one-dimensional vector that consists of row-wise concatenations of beamlet intensities for all fields. Each column of the matrix A is a beamlet kernel, corresponding to the dose distribution achieved by one beamlet with unit intensity. The beamlet kernels are pre-computed based on the CT image of the patient, the treatment machine settings and the beam geometry. In this work, we used the Voxel-based Monte Carlo algorithm (VMC) as our dose calculation engine [25, 26]. The gantry geometry and field angles were based on the typical field setup of the Varian Linac treatment machine.

For an efficient calculation, a convex function is usually used as an objective function in the optimization. If we use the L-2 norm of the difference between the delivered dose and the target dose as the objective function of x ($\phi_1(x)$), the treatment planning problem can now be expressed as:



(a) An actual fluence map obtained using Eclipse planning system on a prostate patient. (b) After taking the discrete gradient as defined in equation (3).

Figure 1: Comparison of an actual fluence map before and after taking discrete gradient.

minimize

$$\phi_1(x) = \sum_i \lambda_i (A_i x - d_i)^T (A_i x - d_i) \quad (2)$$

subject to:

$$x \succeq 0$$

where, the index i denotes different structures; λ_i is the relative importance factor [12, 27–30]; each column of the matrix A_i is the beamlet kernel corresponding to the i 'th structure, and d_i is the prescribed dose.

The main variables used in this paper are summarized in Table 1 for readers' reference.

2.2 Multi-objective optimization including minimization of the segment number

The above optimization problem (2) does not include dose delivery constraints of treatment machines. In practice, the radiation dose is delivered using apertures which are shaped by a multi-leaf collimator. Two constraints are implied. The first is the *uniformity constraint*, i.e. the intensity map of one beam aperture is uniform inside the MLC open area and zero elsewhere if MLC transmission and MLC header scatter can be ignored or compensated for [31, 32]. The second is the *connectivity constraint*, i.e. the non-zero intensity areas of one beam aperture are connected in the direction of MLC leaf pairs. As discussed in the section of introduction, these two constraints are non-convex and not included in the optimization step of the traditional beamlet-based methods, resulting in a large number of beam segments. The number of segments is not optimized in most of the segment-based DAO methods either, since this number is typically pre-fixed before the calculation.

In this paper, we propose an efficient method to optimize the number of segments without compromising the dose distribution. Our algorithm derivation starts from investigating the feature of an intensity map

which can be delivered using a small number of beam segments. Such an intensity map can be considered as a sparse presentation in the space of all deliverable segments, since the total number of possible beam segments obtained by a MLC is extremely large. Another way of understanding the sparsity feature of a desired intensity map can be obtained by considering the uniformity constraint. Mathematically, an actual fluence map is a summation of uniform intensity maps with different shapes. This piece-wise constant function can be easily “sparsified” by taking derivatives. Define a gradient operator as:

$$\nabla_{u,v}x(u,v) = |x_{u,v} - x_{u-1,v}| + |x_{u,v} - x_{u,v-1}| \quad (3)$$

Fig. 1(a) shows an actual fluence map obtained from the Eclipse planning system on a prostate patient. The operator defined in equation (3) is able to sparsify the fluence map, as shown in Fig. 1(b).

The sparsity of the intensity map is well correlated with the corresponding number of segments. The more sparse the optimized intensity map is, the less segments the leaf-sequencing algorithm derives. To enforce the sparsity on the optimized solution and therefore to minimize the number of segments, we include a term of L-1 norm, which is commonly used in compressed sensing algorithms [23,24], as a second objective function in the beamlet-based optimization. Now the multi-objective optimization problem is reformulated as below:

minimize

$$\begin{cases} \phi_1(x) = \sum_i \lambda_i (A_i x - d_i)^T (A_i x - d_i) \\ \phi_2(x) = \sum_{f=1}^{N_f} \sum_{u=2}^{N_u} \sum_{v=2}^{N_v} |\nabla_{u,v}x(u,v,f)| \end{cases} \quad (4)$$

subject to:

$$x \succeq 0$$

where the second objective $\phi_2(x)$ is the L-1 norm of the gradient of the beamlet intensity x as defined in equation (3). x is parameterized by the variables u , v and f . The variables u (v) is the row (column) index of the beam intensity for each field and f is the field index. N_u is the total number of possible MLC leaf positions for each leaf; N_v is the total number of MLC leaves per field; N_f is the number of fields. For simplicity, we assume that each treatment field has a square shape when it is fully open, and N_u and N_v do not change for different fields. Note that $\phi_2(x)$ is actually a total-variation objective function, which is commonly used in many applications to encourage a piece-wise constant solution [33,34].

2.3 Calculation of Pareto frontier

In order to obtain a final solution of the multi-objective optimization problem (5), we choose to first calculate the Pareto frontier and then select the solution which satisfies the clinical acceptance criteria with the smallest

number of segments. The main reason is that some of the clinical goals are non-convex and difficult to be included in the optimization as constraints [35], and visual inspections are used to judge whether a certain plan is clinically acceptable.

The function $\phi_2(x)$ is not linear or quadratic. For an efficient calculation, we reformulate the optimization problem (5) into an equivalent form:

$$\begin{aligned} & \text{minimize} \\ & \begin{cases} \phi_1(x) = \sum_i \lambda_i (A_i x - d_i)^T (A_i x - d_i) \\ \phi_2(x) = e^T t \end{cases} \\ & \text{subject to:} \end{aligned} \tag{5}$$

$$\begin{aligned} x & \succeq 0 \\ Bx - t & \preceq 0 \\ Bx + t & \succeq 0 \end{aligned}$$

where, e is an all-one vector, with a size of $((N_u - 1)N_v N_f + N_u(N_v - 1)N_f)$ -by-1, i.e., $e^T = (1, 1, 1, \dots, 1)$, $e \in \mathbb{R}^{((N_u - 1)N_v N_f + N_u(N_v - 1)N_f) \times 1}$; the vector t is an intermediate variable with the same size as e ; the matrix B is used to calculate the derivatives of x . Specifically,

$$B = \begin{bmatrix} B_h \\ B_v \end{bmatrix} \tag{6}$$

where, B_h is used to calculate the derivatives in the u direction:

$$B_h = \begin{bmatrix} C_1 & & & 0 \\ & C_2 & & \\ & & \ddots & \\ 0 & & & C_{N_v N_f} \end{bmatrix} \tag{7}$$

C_i 's are identical, with a size of $(N_u - 1)$ -by- N_u :

$$C_i = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \tag{8}$$

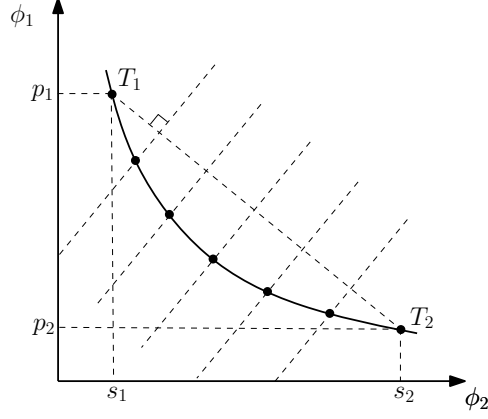


Figure 2: The Pareto frontier of the multi-objective problem. T_1 and T_2 are the anchor points.

B_v is used to calculate the derivatives in the v direction:

$$B_v = \begin{bmatrix} D_1 & & & 0 \\ & D_2 & & \\ & & \ddots & \\ 0 & & & D_{N_f} \end{bmatrix} \quad (9)$$

D_i 's are identical, with a size of $N_u(N_v - 1)$ -by- $N_u N_v$:

$$D_i = \begin{bmatrix} -1 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{bmatrix} \quad (10)$$

On each row, -1 and 1 are separated by $N_u - 1$ zeroes.

2.3.1 Calculation of anchor points

In order to obtain the Pareto frontier, we first fix $\phi_2(x)$ to a small value of s_1 and minimize $\phi_1(x)$ using the following quadratic optimization to obtain an objective value of p_1 :

minimize

$$\phi_1(x) = \sum_i \lambda_i (A_i x - d_i)^T (A_i x - d_i) \quad (11)$$

subject to:

$$\begin{aligned}
x &\succeq 0 \\
Bx - t &\preceq 0 \\
Bx + t &\succeq 0 \\
e^T t &= s_1
\end{aligned}$$

Repeat the optimization using a large $\phi_2(x)$ value of s_2 , and obtain a minimized $\phi_1(x)$ value of p_2 . Thus, we find two anchor points on the Pareto frontier, T_1 and T_2 , as illustrated in Fig. 2.

2.3.2 Calculation of Pareto efficient points between anchor points

In order to calculate the complete Pareto frontier between the two anchor points, one solution is repeating the above optimization using different s values uniformly distributed between s_1 and s_2 . This approach, however, doesn't achieve uniformly distributed data points on the Pareto frontier, as clear in Fig. 2. A better method is used in this work [36]. We minimize the values of $\phi_1(x)$ and $\phi_2(x)$ along lines perpendicular to the line connecting T_1 and T_2 . The function values of $\phi_1(x)$ and $\phi_2(x)$ corresponding to the optimized x define a point on the Pareto frontier. Mathematically, the optimization is changed to be:

minimize

$$\phi_2(x) = e^T t \tag{12}$$

subject to:

$$\begin{aligned}
x &\succeq 0 \\
Bx - t &\preceq 0 \\
Bx + t &\succeq 0 \\
\sum_i \lambda_i (A_i x - d_i)^T (A_i x - d_i) &= g e^T t + h
\end{aligned}$$

where the variable g is the slope of the lines perpendicular to the line $T_1 T_2$, $g = \frac{s_2 - s_1}{p_1 - p_2}$; the variable h is the intercept of these lines. Denote h_1 or h_2 as the intercept of the line passing through T_1 or T_2 , $h_{1,2} = p_{1,2} - g s_{1,2}$. The optimization is repeated for different values of h , which are chosen uniformly between h_1 and h_2 .

Note that the last constraint in the above formulation of optimization defines a non-convex solution set, which makes the problem challenging. Fortunately, it can be verified that this constraint can be changed to be convex without affecting the solution. The optimization becomes a linear programming with linear and quadratic constraints, as shown below:

minimize

$$\phi_2(x) = e^T t \quad (13)$$

subject to:

$$\begin{aligned} x &\succeq 0 \\ Bx - t &\preceq 0 \\ Bx + t &\succeq 0 \\ \sum_i \lambda_i (A_i x - d_i)^T (A_i x - d_i) &\leq g e^T t + h \end{aligned}$$

2.4 Leaf sequencing

The proposed method uses compressed sensing to ensure the sparsity of the optimized solution. However, the obtained beamlet intensity map is not exactly piece-wise constant. Furthermore, the connectivity constraint due to the MLC hardware is not considered in the algorithm derivation. A leaf sequencing algorithm as in beamlet optimization is therefore needed to finally generate deliverable beam segments. Our multi-objective optimization does not post special requirements on the leaf-sequencing step and any existing leaf-sequencing algorithms can be used in combination with the proposed method.

2.5 Evaluation

The proposed algorithm has been tested on a prostate patient. The algorithm was implemented in Matlab, using the MOSEK optimization software package (<http://www.mosek.com>). The anchor points of the Pareto frontier are first calculated using a standard quadratic optimization routine provided in MOSEK with an interior-point optimizer, according to the problem formulation (11). Other Pareto efficient points are calculated using a linear programming with linear and quadratic constraints as shown in (13).

Five fields were used at angles of 35, 110, 180, 250 and 325 degrees, based on a standard clinical protocol for prostate patients. Each field targeted the center of PTV, and contained 20-by-16 beamlets, with a beamlet size of 5mm-by-5mm at the source-to-axis distance (SAD). To save computation, the CT data were downsampled in the dose calculation, and the voxel size was 3.92mm-by-3.92mm-by-2.5mm. The rectum, bladder and femoral heads were included as sensitive structures. All the plans are normalized such that 95% of the PTV volume receives 100% prescribed dose (78Gy).

3 Results

Fig. 3 shows the calculated Pareto frontier of the prostate plan. The number of segments (N_t) corresponding to each Pareto efficient point after applying a leaf-sequencing algorithm is marked in the plot. In general,

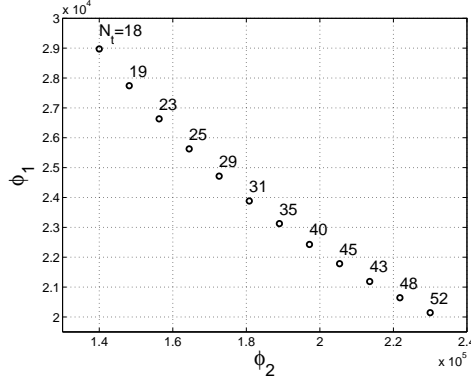


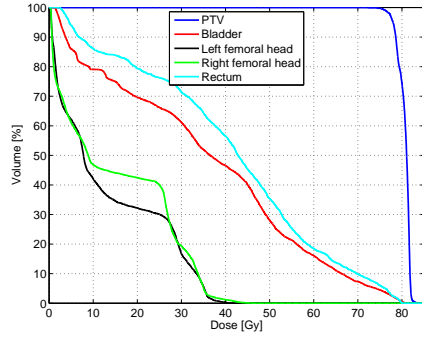
Figure 3: The calculated Pareto frontier of the prostate plan. The derived number of segments (N_t) corresponding to each data point is marked in the plot.

a small (large) $\phi_2(x)$ value achieves a small (large) number of segments, while the dose distribution is degraded (improved), as indicated by the increase (decrease) of $\phi_1(x)$ value. However, since the total-variation objective in our algorithm only implies the uniformity constraint and the connectivity constraint is enforced by the subsequent leaf sequencing, the above relationship is not exactly monotonic. As shown in Fig. 3, in some local areas, a larger $\phi_2(x)$ value achieves a smaller number of segments.

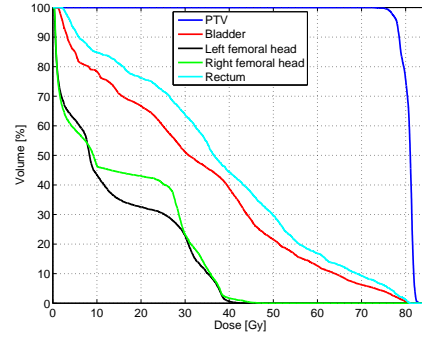
Fig. 4 shows the dose volume histograms (DVH's) of the prostate plans corresponding to every other Pareto efficient point on Fig. 3. As the number of segments increases, the plan performance, especially the avoidance of the organs at risk (OAR's), improves. The improvement slows down when the number of segments reaches a certain level. These plans are evaluated using clinical acceptance criteria and the results are summarized in Table 2. The plan is satisfactory when the segment number is not less than 35, and the result using 35 segments is chosen as the final solution. Using the Eclipse planning system on the same data, the total number of segments is 61. Our method significantly reduces the the number of segments without compromising the clinical performance of the treatment plan. The iso-dose distributions using different numbers of segments are shown in Fig. 5.

4 Discussion and conclusions

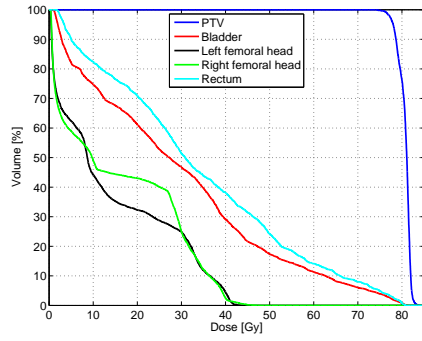
In this paper, an efficient algorithm is proposed to optimize the number of segments without compromising the dose performance in radiation therapy treatment. The derivation is inspired by the feature of sparsity of a desired optimal solution. Using compressed sensing, we include a total-variation term as a second objective function in addition to the traditional objective function in a beamlet-based optimization and reformulate the planning into a multi-objective optimization problem. A method of calculating the Pareto frontier is also designed. Pareto optimal solutions are evaluated using clinical acceptance criteria, and the satisfactory



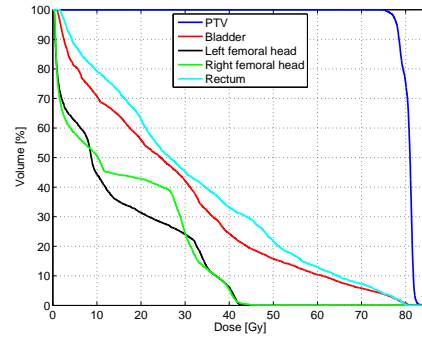
(a) $N_t = 18$



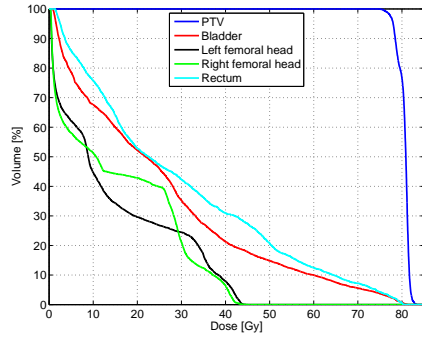
(b) $N_t = 23$



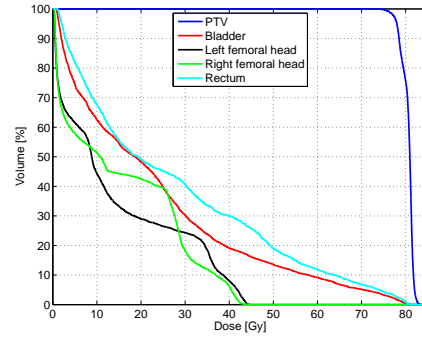
(c) $N_t = 29$



(d) $N_t = 35$

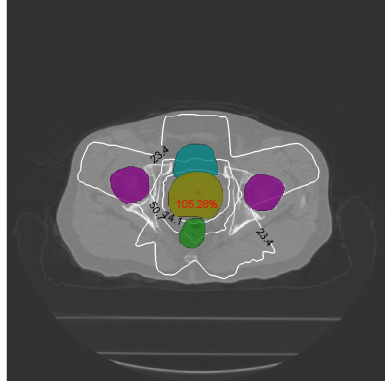


(e) $N_t = 45$

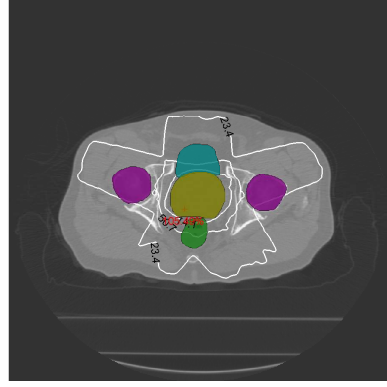


(f) $N_t = 48$

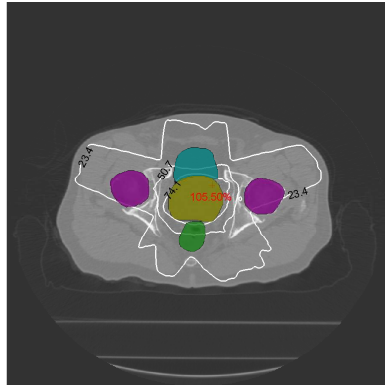
Figure 4: DVH's for the prostate plan using different numbers of segments.



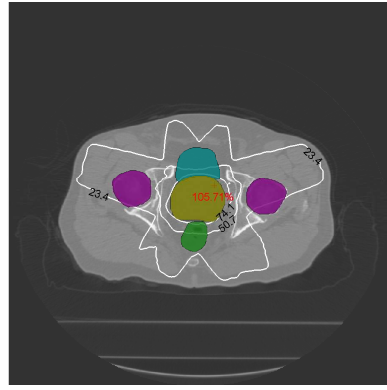
(a) $N_t = 18$



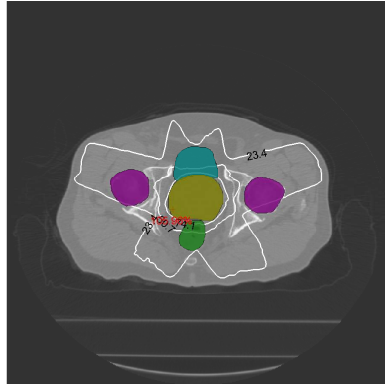
(b) $N_t = 23$



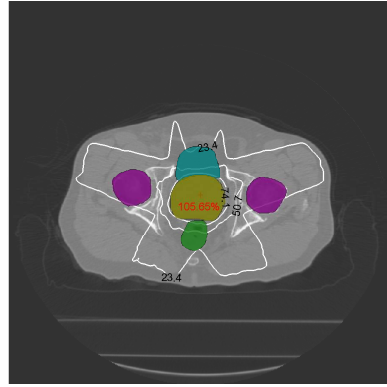
(c) $N_t = 29$



(d) $N_t = 35$



(e) $N_t = 45$



(f) $N_t = 48$

Figure 5: Dose distributions of the prostate plan. The iso-dose lines correspond to 95%, 65% and 30% of the prescribed dose ($78Gy$). The PTV and the sensitive structures (bladder, rectum and femoral heads) are patched using different colors. The hotspots are marked using red crosses.

Table 2: Prostate plan goals and results

Regions	Acceptance criteria	$N_t = 18$	$N_t = 23$	$N_t = 29$	$N_t = 35$	$N_t = 45$	$N_t = 48$
PTV	% vol>78Gy ≥ 95	95.0	95.0	95.0	95.0	95.0	95.0
Rectum	% vol>40Gy ≤ 35	56.5	44.3	38.2	33.3	31.0	30.0
	% vol>65Gy ≤ 17	13.8	12.9	10.7	9.8	9.7	9.4
	vol>79.6Gy $\leq 1cc$	0.50cc	1.42cc	1.27cc	0.54cc	0.81 cc	0.87cc
Bladder	% vol>40Gy ≤ 50	46.5	38.8	29.1	24.3	21.3	19.1
	% vol>65Gy ≤ 25	11.1	9.3	8.1	7.9	7.5	6.9
Femoral heads	% vol>45Gy ≤ 1	0.08	0.30	0.20	0.15	0.03	0
Body	vol>82.7Gy $\leq 1cc$	0.65cc	0.46cc	1.61cc	0.73cc	0.96cc	0.85cc

%vol> x Gy: percentage of the volume that receives more than x Gy dose.

vol> x Gy: size of the volume that receives more than x Gy dose.

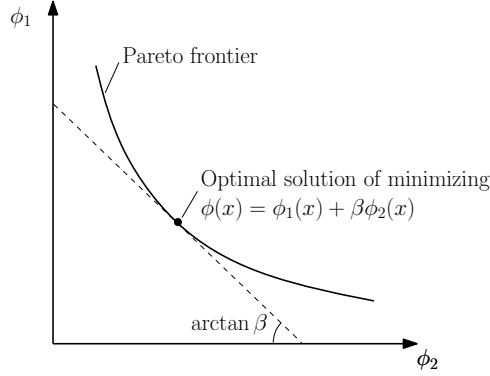


Figure 6: The optimal point on the Pareto frontier if the optimization is solved using a regularization based method.

plan with the smallest number of segment is chosen as the final solution. The algorithm is assessed using a prostate study. The result shows that the proposed greatly reduces the number of segments without compromising the clinical performance of the treatment plan.

Calculation of the Pareto frontier is one solution to a multi-objective optimization problem. Other standard methods can also be used here. For example, we can combine the two objectives and consider the total-variation term as a regularization term with a user-defined penalty weight of β . The optimization problem is converted to a quadratic programming. As shown in Fig. 6, the optimal solution obtained using this method is the Pareto efficient point on the Pareto frontier at which the tangent has a slope of $-\beta$. The multi-objective approach provides a more general solution without introducing the parameter β [35].

The difficulty of optimizing the number of segments in radiation treatment planning arises from the non-convexity of the physical constraints of apertures. Some researchers have proposed sophisticated methods to

find an optimal set of apertures, typically using a formulation of non-convex optimization (more references?) [37]. The computation complexity is increased and a global optimal is not always guaranteed. In our method, we use compressed sensing to encourage the final solution to satisfy the piece-wise constant constraint of an actual fluence map which can be delivered using a small number of segments. The optimization is still convex, and therefore a global optimal solution can always be obtained with a high computation efficiency.

The reduction of the number of segments in a radiation therapy treatment has many indications in clinical practice. First of all, the total treatment time is greatly reduced and therefore the efficiency of the hospital facilities is increased. Second, the reduced treatment time also reduces the patient motion during the treatment and the radiation dose can be delivered more accurately. Third,... As such, the proposed algorithm is very attractive in clinic.

5 Acknowledgement

The authors would like to thank Dr. Alexander Schlaefter for insightful discussion. This project is supported in part by.....

References

- [1] G. A. Ezzell, J. M. Galvin, D. Low, J. R. Palta, I. Rosen, M. B. Sharpe, P. Xia, Y. Xiao, L. Xing, C. X. Yu, I. M. R. T. subcommittee, and A. A. P. M. R. T. committee, "Guidance document on delivery, treatment planning, and clinical implementation of imrt: report of the imrt subcommittee of the aapm radiation therapy committee." *Med Phys*, vol. 30, no. 8, pp. 2089–2115, Aug 2003.
- [2] L. Xing, Y. Wu, Y. Yang, and A. Boyer, *Intensity Modulated Radiation Therapy*. BC Decker Inc.: Hamilton & London, 2005, ch. Physics of intensity modulated radiation therapy, pp. 20–52.
- [3] T. R. Bortfeld, D. L. Kahler, T. J. Waldron, and A. L. Boyer, "X-ray field compensation with multileaf collimators." *Int J Radiat Oncol Biol Phys*, vol. 28, no. 3, pp. 723–730, Feb 1994.
- [4] P. Xia and L. J. Verhey, "Multileaf collimator leaf sequencing algorithm for intensity modulated beams with multiple static segments." *Med Phys*, vol. 25, no. 8, pp. 1424–1434, Aug 1998.
- [5] S. Kamath, S. Sahni, J. Palta, S. Ranka, and J. Li, "Optimal leaf sequencing with elimination of tongue-and-groove underdosage." *Phys Med Biol*, vol. 49, no. 3, pp. N7–19, Feb 2004.
- [6] C. B. Saw, R. C. Siochi, K. M. Ayyangar, W. Zhen, and C. A. Enke, "Leaf sequencing techniques for mlc-based imrt." *Med Dosim*, vol. 26, no. 2, pp. 199–204, 2001.

- [7] H. G. Kuterdem and P. S. Cho, “Leaf sequencing with secondary beam blocking under leaf positioning constraints for continuously modulated radiotherapy beams.” *Med Phys*, vol. 28, no. 6, pp. 894–902, Jun 2001.
- [8] L. Ma, A. L. Boyer, C. M. Ma, and L. Xing, “Synchronizing dynamic multileaf collimators for producing two-dimensional intensity-modulated fields with minimum beam delivery time.” *Int J Radiat Oncol Biol Phys*, vol. 44, no. 5, pp. 1147–1154, Jul 1999.
- [9] M. Alber and F. Nsslin, “Intensity modulated photon beams subject to a minimal surface smoothing constraint.” *Phys Med Biol*, vol. 45, no. 5, pp. N49–N52, May 2000.
- [10] L. Ma, “Smoothing intensity-modulated treatment delivery under hardware constraints.” *Med Phys*, vol. 29, no. 12, pp. 2937–2945, Dec 2002.
- [11] S. V. Spirou, N. Fournier-Bidoz, J. Yang, C. S. Chui, and C. C. Ling, “Smoothing intensity-modulated beam profiles to improve the efficiency of delivery.” *Med Phys*, vol. 28, no. 10, pp. 2105–2112, Oct 2001.
- [12] S. Webb, D. J. Convery, and P. M. Evans, “Inverse planning with constraints to generate smoothed intensity-modulated beams.” *Phys Med Biol*, vol. 43, no. 10, pp. 2785–2794, Oct 1998.
- [13] X. Sun and P. Xia, “A new smoothing procedure to reduce delivery segments for static mlc-based imrt planning.” *Med Phys*, vol. 31, no. 5, pp. 1158–1165, May 2004.
- [14] Y. Xiao, D. Michalski, Y. Censor, and J. M. Galvin, “Inherent smoothness of intensity patterns for intensity modulated radiation therapy generated by simultaneous projection algorithms.” *Phys Med Biol*, vol. 49, no. 14, pp. 3227–3245, Jul 2004.
- [15] M. M. Matuszak, E. W. Larsen, and B. A. Fraass, “Reduction of imrt beam complexity through the use of beam modulation penalties in the objective function.” *Med Phys*, vol. 34, no. 2, pp. 507–520, Feb 2007.
- [16] D. M. Shepard, M. A. Earl, X. A. Li, S. Naqvi, and C. Yu, “Direct aperture optimization: a turnkey solution for step-and-shoot imrt.” *Med Phys*, vol. 29, no. 6, pp. 1007–1018, Jun 2002.
- [17] D. Michalski, Y. Xiao, Y. Censor, and J. M. Galvin, “The dose-volume constraint satisfaction problem for inverse treatment planning with field segments.” *Phys Med Biol*, vol. 49, no. 4, pp. 601–616, Feb 2004.
- [18] C. Cotrutz and L. Xing, “Segment-based dose optimization using a genetic algorithm.” *Phys Med Biol*, vol. 48, no. 18, pp. 2987–2998, Sep 2003.

- [19] B. van Asselen, M. Schwarz, C. van Vliet-Vroegindeweij, J. V. Lebesque, B. J. Mijnheer, and E. M. F. Damen, “Intensity-modulated radiotherapy of breast cancer using direct aperture optimization.” *Radiother Oncol*, vol. 79, no. 2, pp. 162–169, May 2006.
- [20] J. L. Bedford and S. Webb, “Direct-aperture optimization applied to selection of beam orientations in intensity-modulated radiation therapy.” *Phys Med Biol*, vol. 52, no. 2, pp. 479–498, Jan 2007.
- [21] A. M. Bergman, K. Bush, M.-P. Milette, I. A. Popescu, K. Otto, and C. Duzenli, “Direct aperture optimization for imrt using monte carlo generated beamlets.” *Med Phys*, vol. 33, no. 10, pp. 3666–3679, Oct 2006.
- [22] A. Mestrovic, M.-P. Milette, A. Nichol, B. G. Clark, and K. Otto, “Direct aperture optimization for online adaptive radiation therapy.” *Med Phys*, vol. 34, no. 5, pp. 1631–1646, May 2007.
- [23] D. L. Donoho, “Compressed sensing,” *IEEE Trans Information Theory*, vol. 52, pp. 1289–1306, 2006.
- [24] G.-H. Chen, J. Tang, and S. Leng, “Prior image constrained compressed sensing (piccs): a method to accurately reconstruct dynamic ct images from highly undersampled projection data sets.” *Med Phys*, vol. 35, no. 2, pp. 660–663, Feb 2008.
- [25] I. Kawrakow, M. Fippel, and K. Friedrich, “3D electron dose calculation using a Voxel based Monte Carlo algorithm (VMC).” *Med Phys*, vol. 23, no. 4, pp. 445–457, Apr 1996.
- [26] I. Kawrakow, “Improved modeling of multiple scattering in the voxel monte carlo model.” *Med Phys*, vol. 24, no. 4, pp. 505–517, Apr 1997.
- [27] L. Xing, J. G. Li, S. Donaldson, Q. T. Le, and A. L. Boyer, “Optimization of importance factors in inverse planning.” *Phys Med Biol*, vol. 44, no. 10, pp. 2525–2536, Oct 1999.
- [28] S. Breedveld, P. R. M. Storchi, M. Keijzer, and B. J. M. Heijmen, “Fast, multiple optimizations of quadratic dose objective functions in IMRT.” *Phys Med Biol*, vol. 51, no. 14, pp. 3569–3579, Jul 2006.
- [29] T. Bortfeld, “Optimized planning using physical objectives and constraints.” *Semin Radiat Oncol*, vol. 9, no. 1, pp. 20–34, Jan 1999.
- [30] U. Oelfke and T. Bortfeld, “Inverse planning for x-ray rotation therapy: a general solution of the inverse problem.” *Phys Med Biol*, vol. 44, no. 4, pp. 1089–1104, Apr 1999.
- [31] L. Xing and J. G. Li, “Computer verification of fluence map for intensity modulated radiation therapy.” *Med Phys*, vol. 27, no. 9, pp. 2084–2092, Sep 2000.

- [32] Y. Yang and L. Xing, “Incorporating leaf transmission and head scatter corrections into step-and-shoot leaf sequences for imrt.” *Int J Radiat Oncol Biol Phys*, vol. 55, no. 4, pp. 1121–1134, Mar 2003.
- [33] K. T. Block, M. Uecker, and J. Frahm, “Undersampled radial mri with multiple coils. iterative image reconstruction using a total variation constraint.” *Magn Reson Med*, vol. 57, no. 6, pp. 1086–1098, Jun 2007.
- [34] V. Kolehmainen, A. Vanne, S. Siltanen, S. Järvenp, J. P. Kaipio, M. Lassas, and M. Kalke, “Parallelized bayesian inversion for three-dimensional dental x-ray imaging.” *IEEE Trans Med Imaging*, vol. 25, no. 2, pp. 218–228, Feb 2006.
- [35] A. Schlaefter and A. Schweikard, “Stepwise multi-criteria optimization for robotic radiosurgery,” *Med. Phys.*, vol. 35, no. 5, pp. 2094–2103, 2008.
- [36] D. Craft, T. Halabi, and T. Bortfeld, “Exploration of tradeoffs in intensity-modulated radiotherapy,” *Phys. Med. Biol.*, vol. 50, pp. 5857–5868, 2005.
- [37] H. E. Romeijn, R. K. Ahuja, J. F. Dempsey, and A. Kumar, “A column generation approach to radiation therapy treatment planning using aperture modulation,” *SIAM J. on Optimization*, vol. 15, no. 3, pp. 838–862, 2005.